

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:11 ; Search time 55.4461 Seconds
(without alignments)
3839.562 Million cell updates/sec

Title: US-10-797-893-2

Perfect score: 1878

Sequence: 1 MAAGSGTQORREMAAASAA.....FYLDTVSALNFAARKEVIN 370

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1878	100.0	665	1 KF22 HUMAN	Q14807 homo sapien
2	1878	100.0	665	2 AAP35923	Aap35923 homo sapi
3	1604.5	85.4	660	2 Q991C7	Q991C7 mus musculus
4	1232	65.6	639	2 Q6P3R1	Q6P3R1 xenopus tro
5	1232	65.6	639	2 AAHG3896	Aahg3896 xenopus t
6	1202.5	64.0	651	2 Q91869	Q91869 xenopus lae
7	1202.5	64.0	651	2 AAH70549	Aah70549 xenopus l
8	1201.5	64.0	651	2 Q918K0	Q918K0 xenopus lae
9	1197.5	63.8	663	2 Q7ZYL5	Q7zy15 xenopus lae
10	1192.5	63.5	650	2 Q919A8	Q919a8 xenopus lae
11	1191.5	63.4	631	2 Q6GPG0	Q6gpg0 xenopus lae
12	951	50.6	198	2 Q60845	Q60845 homo sapien
13	704	37.5	148	2 Q35232	Q35232 mus musculus
14	623.5	33.2	584	2 Q6L512	Q6l512 oryza sativ
15	623.5	33.2	584	2 AAT39162	Aat39162 oryza sat
16	594.5	31.7	628	2 Q8L7B8	Q8l7b8 arabidopsis
17	594.5	31.7	648	2 Q9LZ88	Q9l288 arabidopsis
18	594.5	31.1	564	2 Q8SQ09	Q8sq09 encephalito
19	578	30.8	912	2 Q7RX60	Q7rx60 neurospora
20	568.5	30.3	821	2 Q7QDS6	Q7qds6 anopheles g
21	563	30.0	838	2 Q862B6	Q862b6 botrytis ci
22	559.5	29.8	784	2 Q961H5	Q961h5 drosophila
23	559	29.8	548	2 Q8N1X8	Q8n1x8 homo sapien
24	555.5	29.6	784	1 KL68 DROME	P46867 drosophila
25	552.5	29.4	898	2 Q86V55	Q86v55 homo sapien
26	552.5	29.4	898	2 Q8NI77	Q8ni77 homo sapien
27	552.5	29.4	898	2 Q9H0F3	Q9h0f3 homo sapien
28	552	29.4	1296	2 Q7RM16	Q7rm16 plasmodium
29	551.5	29.4	562	2 Q8BY99	Q8by99 mus musculus
30	551.5	29.4	571	2 Q8B245	Q8b245 mus musculus
31	551.5	29.4	880	2 Q8BL11	Q8bl11 mus musculus

32 551.5 29.4 886 2 Q91WD7 Q91wd7 mus musculus
33 551.5 29.4 895 2 Q7ZUW9 Q7zuw9 brachydanio
34 550 29.3 703 2 Q9FZ77 Q9fz77 arabidopsis
35 549 29.2 750 2 Q7KSK2 Q7k2 drosophila
36 549 29.2 750 2 AAS65150 Aas65150 drosophil
37 547 29.1 1254 2 Q94463 Q94463 dictyosteli
38 546 29.1 997 2 Q7ZXX2 Q7zxx2 xenopus lae
39 545.5 29.0 389 2 Q7SYZ3 Q7syz3 xenopus lae
40 542.5 28.9 553 2 Q7P8H1 Q7p8h1 anopheles g
41 541 28.8 650 2 Q8DDE9 Q8dde9 xenopus lae
42 536.5 28.6 1058 2 Q9LZU5 Q9lzu5 arabidopsis
43 536 28.5 784 1 KLP6 SCHPO O59751 schizosacch
44 535.5 28.5 677 2 Q9VRK9 Q9vrk9 drosophila
45 533.5 28.4 642 2 Q7QY55 Q7qy55 giardia lam

ALIGNMENTS

RESULT 1

KF22_HUMAN
ID KF22_HUMAN STANDARD; PRT; 665 AA.
AC Q14807; O94814; Q9BT46;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Kinesin-like protein KIF22 (Kinesin-like DNA-binding protein)
DE (Kinesin-like protein 4).
GN Name=KIF22; Synonyms=KNSL4, KID;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RX MEDLINE=96174806; PubMed=8599929;
RA Tokai N., Fujimoto-Nishiyama A., Toyoshima Y., Yonemura S.,
RA Tsukita S., Inoue J., Yamamoto T.,
RT "Kid, a novel kinesin-like DNA binding protein, is localized to
chromosomes and the mitotic spindle."
RL EMBO J. 15:457-467(1996).
SEQUENCE FROM N.A.
RP MEDLINE=99009323; PubMed=9790757;
RX Song J., Murakami H., Yang Z.Q., Koga C., Adati N., Murata T.,
RA Gelling C., Saito-Ohara F., Ikeuchi T., Matsumura M., Itakura K.,
RA Kanazawa I., Sun K., Yokoyama K.K.;
RT "Human genes for KNSL4 and MAZ are located close to one another on
chromosome 16p11.2."
RL Genomics 52:374-377(1998).
SEQUENCE FROM N.A.
RP TISSUE=Lymphocytoblastoma;
RX MEDLINE=99009323; PubMed=9790757;
RA Song J., Murakami H., Yang Z.Q., Koga C., Adati N., Murata T.,
RA Gelling C., Saito-Ohara F., Ikeuchi T., Matsumura M., Itakura K.,
RA Kanazawa I., Sun K., Yokoyama K.K.;
RT "Human genes for KNSL4 and MAZ are located close to one another on
chromosome 16p11.2."
RL Genomics 52:374-377(1998).
SEQUENCE FROM N.A.
RP TISSUE=Brain, and Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmer C.M., Schuler G.D.,
RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:11 ; Search time 76.7254 Seconds
(without alignments)
3839.562 Million cell updates/sec

Title: US-10-797-893-4
Perfect score: 2589
Sequence: 1 MPAAGSTQORREMAASA.....LEAKMLAQAEKENHCPTM 512

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1825181 seqs, 575374646 residues
Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_02:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2577	99.5	665	1 KF22_HUMAN	Q14807 homo sapien
2	2577	99.5	665	2 AAP35923	Aap35923 homo sapi
3	2158.5	83.4	660	2 Q991C7	Q991c7 mus musculus
4	1463	56.5	639	2 Q6P3R1	Q6p3r1 xenopus tro
5	1433	56.5	639	2 AAH63896	AAh63896 xenopus t
6	1438	55.5	651	2 Q91869	Q91869 xenopus lae
7	1438	55.5	651	2 AAH70549	AAh70549 xenopus l
8	1437	55.5	651	2 Q918K0	Q918k0 xenopus lae
9	1428	55.2	663	2 Q7ZVL5	Q7zvl5 xenopus lae
10	1423	55.0	650	2 Q919A8	Q919a8 xenopus lae
11	1422	54.9	631	2 Q6GPG0	Q6gpg0 xenopus lae
12	946	36.5	198	2 Q60845	Q60845 homo sapien
13	704	27.2	148	2 Q35232	Q35232 mus musculus
14	633	24.4	584	2 Q6L512	Q6l512 oryza sativ
15	633	24.4	584	2 AAT39162	Aat39162 oryza sat
16	616	23.8	784	2 Q961H5	Q961h5 drosophila
17	614	23.7	628	2 Q8L7B8	Q8l7b8 arabidopsis
18	612	23.7	664	2 Q9L288	Q9l288 arabidopsis
19	612	23.6	784	1 KL68_DROME	P46867 drosophila
20	611.5	23.6	838	2 Q86ZB6	Q86zb6 botrytis ci
21	608.5	23.5	548	2 Q8SQ09	Q8sq09 neocphalito
22	604	23.3	912	2 Q7RX60	Q7rx60 neurospora
23	603.5	23.3	821	2 Q7QDS6	Q7qds6 anopheles g
24	597.5	23.1	703	2 Q9FZ77	Q9fz77 arabidopsis
25	593	22.9	677	2 Q9VRK9	Q9vrk9 drosophila
26	584.5	22.6	642	2 Q7QY55	Q7qy55 giardia lam
27	580.5	22.4	1394	2 Q7M6Z4	Q7m6z4 mus musculus
28	579.5	22.4	1226	1 KFA4_XENLA	Q91784 xenopus lae
29	579.5	22.4	1226	2 Q6IRM2	Q6irm2 xenopus lae
30	579.5	22.4	1226	2 AAH70854	AAh70854 xenopus l
31	579	22.4	1463	2 Q9GYZ0	Q9gyz0 stronglyloce

32	578.5	22.3	1058	2 Q9LZU5	Q9lzu5 arabidopsis
33	577	22.3	571	2 Q8B245	Q8bz45 mus musculus
34	577	22.3	880	2 Q8BL11	Q8bl11 mus musculus
35	577	22.3	886	2 Q91WD7	Q91wd7 mus musculus
36	575.5	22.2	1231	2 Q6DIN5	Q6din5 xenopus tro
37	574.5	22.2	787	2 Q6PKB2	Q6pkb2 homo sapien
38	574.5	22.2	787	2 AAH03664	AAh03664 homo sapi
39	574.5	22.2	1127	2 Q86TN3	Q86tn3 homo sapien
40	574.5	22.2	1232	1 KF4A_HUMAN	Q95239 homo sapien
41	574	22.2	562	2 Q8BY99	Q8by99 mus musculus
42	572	22.1	895	2 Q7ZUW9	Q7zuw9 brachydanio
43	570	22.0	699	2 Q98T11	Q98t11 xenopus lae
44	569.5	22.0	915	2 Q6GR48	Q6gr48 xenopus lae
45	568.5	22.0	671	2 Q86XX7	Q86xx7 homo sapien

ALIGNMENTS

RESULT 1
KF22_HUMAN
ID KF22_HUMAN STANDARD; PRT; 665 AA.
AC Q14807; Q94814; Q9BT46;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Kinesin-like protein KIF22 (Kinesin-like DNA-binding protein)
DE (Kinesin-like protein 4).
DE Name=KIF22; Synonyms=KNSL4, KID;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96174806; PubMed=8599929;
RA Tokai N., Fujimoto-Nishiyama A., Toyoshima Y., Yonemura S.,
TSukita S., Inoue J., Yamamoto T.;
RT "Kid, a novel kinesin-like DNA binding protein, is localized to
RT chromosomes and the mitotic spindle."
RL EMBO J. 15:457-467(1996).
RN [2]
RP SEQUENCE FROM N.A.
RT TISSUE=Lymphocytoblastoma;
RX MEDLINE=99009123; PubMed=9790757;
RA Song J., Murakami H., Yang Z.Q., Koga C., Adati N., Murata T.,
RA Gelling C., Saito-Ohara F., Ikeuchi T., Matsumura M., Itakura K.,
RA Kanazawa I., Sun K., Yokoyama K.K.;
RT "Human genes for KNSL4 and MAZ are located close to one another on
RL chromosome 16p11.2."
RN [3]
RP SEQUENCE FROM N.A.
RT TISSUE=Brain, and Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.M., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

QY 183 EKVLDLDPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASNRRTVGATRLNQ 242
 Db 182 EKVLDLDPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASNRRTVGATRLNQ 241
 QY 243 RSRSHAVLVKVDQRELERAPFQREGKLYLIDLAGSDNRRTGNKGLRKESGAINTSL 302
 Db 242 RSRSHAVLVKVDQRELERAPFQREGKLYLIDLAGSDNRRTGNKGLRKESGAINTSL 301
 QY 303 FVLGKVVADALNOGLPRVYPRYRDSKLTLLQDSLAGSIIILIANIAPERFYLTIVSALNF 362
 Db 302 FVLGKVVADALNOGLPRVYPRYRDSKLTLLQDSLAGSIIILIANIAPERFYLTIVSALNF 361
 QY 363 AARSKEVINPFTNESLOPHALGPVKLSQKELLGPPPEAKRGPPEBEEIGSPPEMAAPAS 422
 Db 362 AARSKEVINPFTNESLOPHALGPVKLSQKELLGPPPEAKRGPPEBEEIGSPPEMAAPAS 421
 QY 423 ASQKLSPLQKLSMDPAMLERLLSLDRLLASQSQGAPLLSTPKRERMVLMKTVEEKDLE 482
 Db 422 ASQKLSPLQKLSMDPAMLERLLSLDRLLASQSQGAPLLSTPKRERMVLMKTVEEKDLE 481
 QY 483 IERLKTQKLEAKMLAQKAEKENHCPTM 512
 Db 482 IERLKTQKLEAKMLAQKAEKENHCPTM 511

RESULT 3
 Q99LC7
 ID Q99LC7 PRELIMINARY; PRT; 660 AA.
 AC Q99LC7
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Kinesin family member 22.
 GN Names-Kif22;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Heiton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalka U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC003427; AA03427.1; -
 DR HSSP; P33173; 1155.
 DR MGD; MGI:109233; Kif22.
 DR GO; GO:0000785; C:chromatin; IDA.
 DR GO; GO:0005819; C:spindle; IDA.
 DR InterPro; IPR003583; HHH_1.

DR InterPro; IPR01752; kinesin motor.
 DR InterPro; IPR010994; Ruva_2_Like.
 DR Pfam; PF00225; Kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00278; HHH1; 2.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
 DR PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.
 KW ATP-binding; Microtubule; Motor protein.
 SQ SEQUENCE 660 AA; 73189 MW; AA8B99477BC49B3C CRC64;

Query Match 83.4%; Score 2158.5; DB 2; Length 660;
 Best Local Similarity 84.8%; Pred. No. 3.8e-132;
 Matches 431; Conservative 32; Mismatches 36; Indels 9; Gaps 4;

QY 8 TCQRREMAAASAAATSGAGRCRLSKTGATRRPPARVAVRLRPVDCGTAGASDPVCV 67
 Db 7 TCQRREMAAATSGAGRCRLSKTGATRRPPARVAVRLRPVDCGTAGASDPVCV 61
 QY 68 RGMDSCLSIATNRNHNQETLKVQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLAY 127
 Db 62 RAIDSCSLELVANWKKYQETLKVQDFAFYGEKSTQQEVYVGVQPIRLHLLLEGQNASVLAY 121
 QY 128 GPTGAGKTHMLGSPQPGVIPPALMDLLQTLTREBGAEGRPWALSVTMSYLEIYQEKVLD 187
 Db 122 GPTGAGKTHMLGSPQPGVIPPALMDLLQTLTREBGAEGRPWALSVTMSYLEIYQEKVLD 181
 QY 188 LLDPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASNRRTVGATRLNQSSRS 247
 Db 182 LLDPASGDLVIREDCRGNILIPGLTQKPISSPADFERHFLPASNRRTVGATRLNQSSRS 241
 QY 248 HAVLLVKVDQRELERAPFQREGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGK 307
 Db 242 HAVLLVKVDQRELERAPFQREGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGK 301
 QY 308 VVDALNOGLPRVYPRYRDSKLTLLQDSLAGSIIILIANIAPERFYLTIVSALNFARSK 367
 Db 302 VVDALNOGLPRVYPRYRDSKLTLLQDSLAGSIIILIANIAPERFYLTIVSALNFARSK 361
 QY 368 EVINRPTNESLOPHALGPVKLSQKELLGPPPEAKRGPPEBEEIGSPPEMAAPASASOKL 427
 Db 362 EVINRPTNESLOPHALGPVKLSQKELLGPPPEAKRGPPEBEEIGSPPEMAAPASASOKL 421
 QY 428 SPLQKLSMDPAMLERLLSLDRLLASQSQGAPLLSTPKRERMVLMKTVEEKDLETERLK 487
 Db 422 SPLQKLSMDPAMLERLLSLDRLLASQSQGAPLLSTPKRERMVLMKTVEEKDLETERLK 481
 QY 488 TKQKLEAKMLAQKAEKENHCPTM 512
 Db 482 TKQKLEAKMLAQKAEKENHCPTM 508

RESULT 4
 Q6P3R1
 ID Q6P3R1 PRELIMINARY; PRT; 639 AA.
 AC Q6P3R1
 DT 05-JUL-2004 (TReMBLrel. 27, Created)
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
 DE Hypothetical protein MGC75575.
 GN Name=MGC75575;
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8364;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

```

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalilus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [12]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC063896; AAH63896.1; -.
DR InterPro; IPR001752; kinesin motor.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1.
DR ATP-binding; Hypothetical protein; Microtubule; Motor protein.
KW ATP-binding; Hypothetical protein; Microtubule; Motor protein.
SQ SEQUENCE 639 AA; 71390 MW; F5664F9020A6A1AA CRC64;

Query Match 56.5%; Score 1463; DB 2; Length 639;
Best Local Similarity 60.3%; Pred. No. 7.8e-87;
Matches 292; Conservative 79; Mismatches 103; Indels 10; Gaps 4;

QY 30 RLSKIGATRRPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLFIANWRNHQETLKY 89
DB 4 RVSLDQHKHKSARVRVAVRLRPFYMEKEDEKAPACVGLGDSQSLFIANWRNHQETLKY 63

QY 90 QPDAFYGERSTQDIYAGSVQPIRLHLEGQNASVLAQPTGAGKTHMLGSPQGVIP 149
DB 64 QPDAFYGDSATQREIYAGSVCHILPHLLIGQNASVLAQPTGAGKTHMLGSPQGVIP 123

QY 150 RALMDLLQLTREE--GAEGRPWALSVMYSLEYIQEKVLDLDPASGDLVIREDCRGNIL 207
DB 124 RAVRDLLQMTTRAAGGPENENWTYITMSYVEIQEKVMDLLEPKNKLPIREDKDHIL 183

QY 208 IGPLSQKPISSFADPERHFLPASNRITVGATRLNQRSSRSRSHAVLLKVDQERLAPRQR 267
DB 184 IGVVTKTINSFGDFDEHFI PASQNRITVASTKLNDRSSRSRSHAVLLIKVQSQVSPFRL 243

QY 268 EGYLYLDLAGSEDRNRRTGNKGLRKESGAINTSFLVLGVVDALNQLPRVPYRDSKLT 327
DB 244 TGLKYLIDLAGSEDRNRRTGNKGLRKESGAINTSFLVLGVVDALNQLPRVPYRDSKLT 303

QY 328 RLQDSLGGSAHSILIANIAPERFRFYDTVSALNFAARKEVINRPFTHESLOPHALGPV 387
DB 304 RLQDSLGGTAHSVMIANIAPERFYDTLTALNFAAKSQIINKPFSQETTSQIAALPA 363

QY 388 KLSQKELLGPPEAKRGPPEEIGSPPEPMAAPASAKSLPQK--LSSMDPAMLERLL 445
DB 364 MKRPRE-----EATAAGSQKQKSKTDSTESSPNTSMDSAKKLNLAALDPVVERLL 418

QY 446 SLDRLLASQGSQGLAPLSTPKRERMLVMTKVEKDLIERLTKYKQKLENAKQAE-E 504
DB 419 KLDKILTEKGMKEAQLSTPKRERMLVMTKVEKDLIERLTKYKQKLENAKQAEARLE 478

QY 505 KENH 508
DB 479 KSTN 482

```

```

RESULT 5
AAH63896 PRELIMINARY; PRT; 639 AA.
AC AAH63896;
DT 25-MAR-2004 (TrEMBLrel. 27, Created)
DT 25-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 25-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein MGC75575.
GN MGC75575.
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Diatchenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalilus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC063896; AAH63896.1; -.
KW Hypothetical protein.
SQ SEQUENCE 639 AA; 71390 MW; F5664F9020A6A1AA CRC64;

Query Match 56.5%; Score 1463; DB 2; Length 639;
Best Local Similarity 60.3%; Pred. No. 7.8e-87;
Matches 292; Conservative 79; Mismatches 103; Indels 10; Gaps 4;

QY 30 RLSKIGATRRPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLFIANWRNHQETLKY 89
DB 4 RVSLDQHKHKSARVRVAVRLRPFYMEKEDEKAPACVGLGDSQSLFIANWRNHQETLKY 63

QY 90 QPDAFYGERSTQDIYAGSVQPIRLHLEGQNASVLAQPTGAGKTHMLGSPQGVIP 149
DB 64 QPDAFYGDSATQREIYAGSVCHILPHLLIGQNASVLAQPTGAGKTHMLGSPQGVIP 123

QY 150 RALMDLLQLTREE--GAEGRPWALSVMYSLEYIQEKVLDLDPASGDLVIREDCRGNIL 207
DB 124 RAVRDLLQMTTRAAGGPENENWTYITMSYVEIQEKVMDLLEPKNKLPIREDKDHIL 183

QY 208 IGPLSQKPISSFADPERHFLPASNRITVGATRLNQRSSRSRSHAVLLKVDQERLAPRQR 267
DB 184 IGVVTKTINSFGDFDEHFI PASQNRITVASTKLNDRSSRSRSHAVLLIKVQSQVSPFRL 243

QY 268 EGYLYLDLAGSEDRNRRTGNKGLRKESGAINTSFLVLGVVDALNQLPRVPYRDSKLT 327
DB 244 TGLKYLIDLAGSEDRNRRTGNKGLRKESGAINTSFLVLGVVDALNQLPRVPYRDSKLT 303

QY 328 RLQDSLGGSAHSILIANIAPERFRFYDTVSALNFAARKEVINRPFTHESLOPHALGPV 387
DB 304 RLQDSLGGTAHSVMIANIAPERFYDTLTALNFAAKSQIINKPFSQETTSQIAALPA 363

QY 388 KLSQKELLGPPEAKRGPPEEIGSPPEPMAAPASAKSLPQK--LSSMDPAMLERLL 445
DB 364 MKRPRE-----EATAAGSQKQKSKTDSTESSPNTSMDSAKKLNLAALDPVVERLL 418

QY 446 SLDRLLASQGSQGLAPLSTPKRERMLVMTKVEKDLIERLTKYKQKLENAKQAE-E 504
DB 419 KLDKILTEKGMKEAQLSTPKRERMLVMTKVEKDLIERLTKYKQKLENAKQAEARLE 478

QY 505 KENH 508
DB 479 KSTN 482

```

Db 304 RLQDSGGTAHSMVIANIAPEQYFDLTALNFAAKSQIINKPFSQETTSIAALPA 363
 QY 388 KLSQKELLGPPKARGPEEIGSPPEWAAASQKSLPQK--LSSMDPAMLERLL 445
 Db 364 MKPRE-----EAEIAGSQKRSKTDSTESPTMSDASRRKUNLAALDPFVVERLL 418
 QY 446 SLDRLLASQSQGAPLLSTPKRMRVLMKVTBCKLEIERLTKTKQKLEAKMLAKAE-E 504
 Db 419 KLDKILTEKMGKQAQLLSTPKRMRVLMKVTBCKLEIERLTKTKQKLEAKMLAKAE-E 478
 QY 505 KENH 508
 Db 479 KSTN 482

RESULT 6
 Q91869 PRELIMINARY; PRT; 651 AA.

AC Q91869; 651 AA.
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Kinesin (Hypothetical protein) (Chromokinesin Xkid).
 GN Nameskid;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20419289; PubMed=10966105;
 RA Antonio C., Ferby I., Wilhelm H., Jones M., Karsenti E., Nebreda A.R.,
 RA Vernos I.,
 RT "Xkid, a chromokinesin required for chromosome alignment on the
 metapase plate.";
 RL Cell 102:425-435(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 initiative.";
 RL Dev. Dyn. 225:384-391(2002).
 RN [4]
 RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;
 RA Klein S., Strausberg R.;
 RA Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blastula;
 RX MEDLINE=99340303; PubMed=10411507;
 RA Zou H., McGarry T.J., Bernal T., Kirschner M.W.;
 RT "Identification of a vertebrate sister-chromatid separation inhibitor
 involved in transformation and tumorigenesis.";
 RL Science 285:418-422(1999).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blastula;
 RX MEDLINE=20419288; PubMed=10966104;
 RA Funabiki H., Murray A.W.;
 RT "The Xenopus chromokinesin Xkid is essential for metaphase chromosome
 alignment and must be degraded to allow anaphase chromosome
 movement.";
 RL Cell 102:411-424(2000).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blastula;
 RA McGarry T.J., Bernal T., Funabiki H., Kirschner M.W.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ249840; CAB71798.1; -;
 DR EMBL; BC070549; AAH70549.1; -;
 DR EMBL; AF267849; AAF82563.1; -;
 DR HSP; P33173; I155.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0005875; C:microtubule associated complex; IEA.
 DR GO; GO:0005924; F:ATP binding; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003774; F:motor activity; IEA.
 DR InterPro; IPR010936; F:Motor activity; IEA.
 DR InterPro; IPR001583; HHH 1.
 DR InterPro; IPR001752; kinesin_motor.
 DR Pfam; PF00225; Kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00278; HHH1; 2.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
 DR PROSITE; PS00677; KINESIN MOTOR DOMAIN2; 1.
 KW ATP-binding; Hypothetical protein; Microtubule; Motor protein.
 SQ SEQUENCE 651 AA; 73065 MW; C2ABA561C4C53C13 CRC64;
 Query Match 55.5%; Score 1438; DB 2; Length 651;
 Best Local Similarity 60.1%; Pred. No. 3.4e-85;
 Matches 292; Conservative 80; Mismatches 100; Indels 14; Gaps 7;
 QY 30 RLKIGATRRPPPARVRVAVRLRPFDVGTAGA--SDPPCVRGMDSCSLEIANWENHOETLK 88
 Db 17 RVSLMDQHKSSCARVRVAVRLRPYNDKDEAKATTCVVRGLDSQSLEIVNWRNQLTQM 76
 QY 89 YOFDAFYGERSTQQDIYAGSVQPIRLHLEQGNASVLAIGPTGAGKTHMLGSPQGV 148
 Db 77 YQDAFYGDSASQREIYGVSVCHPLHLLIGQNASVFAVPTGAGKTHMLGNPQGV 136
 QY 149 PALMDLLQLTREGA--EGRPWALSVTVSYLEIYQEKVLDLDPASGLVIREDCRNI 206
 Db 137 PRAVRDLQMSRTAASAPENENNTYTNNVYVEIYQEKVMDLLEPKNKDLPFIREDKHNI 196
 QY 207 LIPGLSQPISSFADFPERHFLPASNRRTVGCATLNORSRSRSHAVLLVKVDQERLAPFR 266
 Db 197 LIPGVTKMINSFADFDEHFIPASQNRVTASTKJNDRSSRSHAVLLIKVKQSQVVPFR 256
 QY 267 REGKLYLDLAGSEDNRRGTGNKGLRKESGAINTSFLVGVKVDALNQGILPRVYRDSKL 326
 Db 257 LTGKLYLDLAGSEDNRRGTGNQIRLKESGAINSSFLTSKVVDALNQGILPRVYRDSKL 316
 QY 327 TRLLQDSLGGSASHILIANIAPERFYLDTVSALNFAARSKEVINRPFNTNESLQPHALGP 386
 Db 317 TRLLQDSLGGSASHVMTNIAPEQTYFDLTALNFAAKSQIINKPFSQETTSIAALPA 376

QY	387	VKLSQKE---	LLGPPEAKRARGPEEEIGSPPEMAAPASOKLSPLQKLSMDPAMLER	443
Db	377	MKRPREETGHTAGSGKRRKSKNDSTE--SSPN--SSMDTAGKQKLN----	LATLDPAAVVER	429
QY	444	LLSLDLRLAASQGSQAPLLSTPKRERMVMTVEKDLERLTKKOKELBAKMLAQAE	503	
Db	430	LLKLDKILTEKGKKKQALLSTPKRERMALLKKWESQMEIERLKEKOKELQKAMEAEAR	489	
QY	504	-EKENH 508		
Db	490	LEKSN 495		
RESULT 7				
ID	AAH70549	PRELIMINARY;	PRT;	651 AA.
AC	AAH70549;			
DT	13-MAY-2004	(TREMBlrel. 27, Created)		
DT	13-MAY-2004	(TREMBlrel. 27, Last sequence update)		
DT	13-MAY-2004	(TREMBlrel. 27, Last annotation update)		
DE	Hypothetical protein.			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;			
OC	Xenopodinae; Xenopus; Xenopus.			
OX	NCBI_TaxID=8355;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Embryo;			
RX	MEDLINE=22341132; PubMed=12454917;			
RA	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,			
RA	Richardson P.;			
RT	"Genetic and genomic tools for Xenopus research: The NIH Xenopus			
RT	initiative.";			
RL	Dev. Dyn. 225:384-391(2002).			
RP	[2]			
RN	SEQUENCE FROM N.A.			
RC	TISSUE=Embryo;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,			
RA	Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickinson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,			
RA	Krzyszinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E.,			
RA	Jones S.J., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RP	[3]			
RN	SEQUENCE FROM N.A.			
RC	TISSUE=Embryo;			
RA	Klein S., Strausberg R.;			
RL	Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC070549; AAH70549.1;			
KW	Hypothetical protein.			
SEQUENCE	651 AA;	C2ABA561C4C53C13	CRC64;	
Query Match	55.5%;	Score 1438;	DB 2;	Length 651;
Best Local Similarity	60.1%;	Pred. No. 3.4e-85;		
Matches 292;	Conservative	80;	Mismatches 100;	Indels 14;
				Gaps 7;

QY	30	RLSKIGATRRPPPPARVRVAVLRPFVDGTAGA--SDPPCVRGMDSCSLEIATNRNHOETLK	88	
Db	17	RVSMLDQHKSSCARVRVAVLRPFYMDKDEAKATTCVVRGLDSQSLEIATNRNHOETWQ	76	
QY	89	YQFAPFYGERSTOODIYAGSVQPIRLRHLLEGQNASVLAVGPTGAGKTHTWLMSPEQPGVI	148	
Db	77	YQFAPFYGDSASQREIYMGSVCHILPHLLIGQNASVFAVPTGAGKTHTWLGNPNQPGVI	136	
QY	149	PRALMDLIQITFREGA--EGRPWALSVMTSYLIEYQEKVLDLDPASGDLVIREDCRNI	206	
Db	137	PRAVRDLQLMSRTAASAPENENWTYTTINMSYVEIYQEKVMDLLEPKNKDLPFIREDKHNI	196	
QY	207	LIPGLSQKPISSPADFERHFELPASRNTVGCATNLNORSSRSHAVLLVKVDORERLAPFRQ	266	
Db	197	LIPGVYQKRVINSFADFDEHFIPASQNRVTVASTKLNDRSSRSHAVLLVKVKSQVVPFRQ	256	
QY	267	REGKLYLDLAGSDNRRTGNKGLRLKESGAINTSFLVGLKVVDALNQGLPRVPYRDSKL	326	
Db	257	LTGKLYLDLAGSDNRRTGNQGLRLKESGAINSSLFTLSKVVDALNQGLPRIPYRDSKL	316	
QY	327	TRLLODSIGGSAHSILIANIAPERRFYLDITVSALNFAARSKVINRPFNTNESIQHALGP	386	
Db	317	TRLLODSIGGSAHSVMTIATPEQTYFYFDLTALNFAAKSQIINKFSETTQTVVQPA	376	
QY	387	VKLSQKE--LLGPPEAKRARGPEEEIGSPPEMAAPASOKLSPLQKLSMDPAMLER	443	
Db	377	MKRPREETGHTAGSQKRRKSKNDSTE--SSPN--SSMDTAGKQKLN----	429	
QY	444	LLSLDLRLAASQGSQAPLLSTPKRERMVMTVEKDLERLTKKOKELBAKMLAQAE	503	
Db	430	LLKLDKILTEKGKKKQALLSTPKRERMALLKKWESQMEIERLKEKOKELQKAMEAEAR	489	
QY	504	-EKENH 508		
Db	490	LEKSN 495		
RESULT 8				
ID	Q918K0	PRELIMINARY;	PRT;	651 AA.
AC	Q918K0;			
DT	01-OCT-2000	(TREMBlrel. 15, Created)		
DT	01-OCT-2000	(TREMBlrel. 15, Last sequence update)		
DT	01-MAR-2004	(TREMBlrel. 26, Last annotation update)		
DE	Chromokinesin Xkid.			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;			
OC	Xenopodinae; Xenopus.			
OX	NCBI_TaxID=8355;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	MEDLINE=20419288; PubMed=10966104;			
RX	Funabiki H., Murray A.W.;			
RA	"The Xenopus chromokinesin Xkid is essential for metaphase chromosome			
RT	alignment and must be degraded to allow anaphase chromosome			
RT	movement.";			
RL	Cell 102:411-424(2000).			
DR	EMBL; AF267850; AAF82564.1;			
DR	HSSP; P33173; 115S.			
DR	GO; GO:0005622; C:intracellular; IEA.			
DR	GO; GO:0005875; C:microtubule associated complex; IEA.			
DR	GO; GO:0005524; F:ATP binding; IEA.			
DR	GO; GO:0003677; F:DNA binding; IEA.			
DR	GO; GO:0003774; F:motor activity; IEA.			
DR	InterPro; IPR010996; DNaPol_B_N-like.			
DR	InterPro; IPR003583; HHH 1.			
DR	InterPro; IPR001752; kinesin_motor.			
DR	Pfam; PF00225; Kinesin; 1.			
DR	PRINTS; PRO0380; KINESINHEAVY.			
DR	SMART; SM00278; HHh1; 2.			
DR	SMART; SM00129; Kisc; 1.			
DR	PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.			

Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00278; Hhh1; 2.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN MOTOR_DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.
FT NON_TER
SQ SEQUENCE 663 AA; 74538 MW; 3B98B5F3C12C9C22 CRC64;

```

Query Match          55.2%; Score 1428; DB 2; Length 663;
Best Local Similarity 59.4%; Pred. No. 1.6e-84;
Matches 287; Conservative 82; Mismatches 106; Indels 8; Gaps 4
```

Qy	30	RLSKIGATRRPPPARVAVRLRPFVDGTAGASDP-PCVRGMDSCSLSEIANRNHQTGLK	88
Dd	30	RVSMLDQHKKSSCARVAVLRLPVMEKEEDKVPTACVRGLDLSHLEIVWRNQLETWQ	89
Qy	89	YQPDATFGERSQTQDIIVAGSVQPIRLHLLEGONASVLAYGPTGAGKTHTMLSGPEOPGV	148
Dd	90	YQPDATFYGDSASOREIYMGSVCHILPHLLIQNASVFAYGPTGAGKTHTMLGNPQPGVI	149
Qy	149	PRALMDLQLTR--EEGAEGRPWALSVTMSYLEIYOEKVLDLDDPASGDLVIREDCRNI	206
Dd	150	PRAVRELLQMTMAASAPENENWTYTITMSVEIYQEKVMDLLEPKNDLPREDKDHNI	209
Qy	207	LIPGLSKPISSFADFERHFPLPSNRNRTVGATRLNQRSRSRHAVLLVKVDQERLERLAPFRQ	266
Dd	210	LIPGVTLTKINSFCDFDEHFTPASQNRTVASTKLNDRSRSHAVLLIKVKSQQVAFPRQ	269
Qy	267	REGKLYLIDLAGSDNRRTNGKLRLESQAINTSLFVLGKVVADALNOGLPRVPYRDSKL	326
Dd	270	LIGKLYLIDLAGSDNRRTNGQIRLESQAINSFLTSLKVVDALNOGLPRIPYRDSKL	329
Qy	327	TRLLODSLGSASHILLANTAPERFPYDLTVSALNFAARSEVINRPNTPNESLQHPHALGP	386
Dd	330	TRLLODSLGSASHVMITNTAPEQTYFDLTALTALFAAKSKQINKFPSRETQTVAQPA	389
Qy	387	VKLSOKELLGPPEAKRKARGPEEBIGSPPEPWAA PASAQSLSPLOKLSMMDPAMLERLLS	446
Dd	390	MKPREEAEATTSRQRKSKYDTSTESPNSMESTGRKLN----LASLDASVVERLLK	445
Qy	447	LDRLLASQGSQGAPLLSTPKRERMVLMKTVEKDLEIERLTKTKOKELEAKMLAQAB-EK	505
Dd	446	LDKILTCKGKKEAQLLSTPKRERMALLKWEESQWEIERLEKEKELEQAMEAEARLEK	505
Qy	506	ENH 508	
Dd	506	SNN 508	

RESULT 10
Q919A8 PRELIMINARY; PRT; 650 AA.

ID	Q919A8	PRELIMINARY;	PRT;	650 AA.
AC	Q919A8;			
DT	01-OCT-2000 (TrEMBLrel. 15, Created)			
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE	Kinesin.			
GN	Name=kid;			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;			
OC	Xenopodinae; Xenopus.			
OX	NCBI_Taxid=8355;			
RN	[1]			
RX	SEQUENCE FROM N.A.			
RA	MEDLINE=20419289; Pubmed=10966105;			
RA	Antonio C., Ferby I., Wilhelm H., Jones M., Karsenti E., Nebreda A.R.,			
RA	Vernos I.;			
RT	"Xkid, a chromokinesin required for chromosome alignment on the			
RL	metaphase plate.";			
RL	Cell 102:425-435(2000).			

Fri Nov 12 12:26:18 2004

```
DR EMBL; AJ249841; CAB71799.1; -.
DR HSSP; P33173; 115S.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005875; C:mitochondrion associated complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0010996; DNAPH1_B_N-like.
DR InterPro; IPR003583; HH1_1_N-like.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; Hh1; 2.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1.
DR ATP-binding; Microtubule; Motor protein.
KW SEQUENCE 650 AA; 73027 MW; EA415BC4B9B0F4A CRC64;
Query Match 55.0%; Score 1423; DB 2; Length 650;
Best Local Similarity 59.4%; Pred. No. 3.2e-84;
Matches 287; Conservative 81; Mismatches 107; Indels 8; Gaps 4;
QY 30 RLKLGATRRPPAPRVAVRLRPFDGTAGASDP-PCVVRGMDSCSLFIANRNHOETLK 88
Db 17 RASMLDQHKKSCARVAVRLRPYEEKEEDKVTACVRLGDSLSHLSIVNWRQLETMQ 76
QY 89 YQDFAFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAYGPTGAGTKHTMLGSPQPGVI 148
Db 77 YQDFAFYGSASQREIYMGVCHILPHLLIGNASVFAYGPTGAGTKHTMLGNPDQPGVI 136
QY 149 PRALMDLQLTR--EAGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 206
Db 137 PRVRELLQWTRMAASAPENENWTHITMSYVEIYQEKVMDLLEPKNDLPREDKDHNI 196
QY 207 LIPCLSQKDISSFADPERHFLPASNRRTVGNRLNQRSSSHAVLLVKVDORERLAPRQ 266
Db 197 LIPGVTLKINSFGDFHIFPASQNRVASTKLNDRSSSHAVLLIKVQSQVAPRQ 256
QY 267 REGKYLIDLAGSEDRNRRTGNKGLRKESGAINSLFVLGKVVVDALNOGLPRVPRDLSKL 326
Db 257 LICKLYLIDLAGSEDRNRRTGNQIRLKESGAINSLFTLSKVVVDALNOGLPRIPYRDSKL 316
QY 327 TRLLQSLGSAHSIIANIPERRYLDTVSANLFAEKSKEVINRPFNTNLSLOPHALGP 386
Db 317 TRLLQSLGSAHSVMTNIAPEQTYFTLTALNFAAKSKQIINKPFRRTITQVAPQ 376
QY 387 VKLSQKELLGPPAKRGPEEEIGSPPEMAAPASOKLSPLOKLSMDPAMLERLLS 446
Db 377 MKPREAEATTSRORKSKTSTESSPNSSMESTGKLN---LASLDSAVVERLLK 432
QY 447 LDRLLASQSGQAGPLSTPKRERVLMTKTVVEKDLIERLKTKYKQKLEAKMLAQKAE-EK 505
Db 433 LDKILTCKGKKEAQLSTPKRERVALLKKWEESQMBIERLKEKQKLEQKAEAEARLEK 492
QY 506 ENH 508
Db 493 SNN 495
RESULT 11
Q6GPGO PRELIMINARY; PRT; 631 AA.
AC Q6GPGO;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
```

```
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Heiton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RA and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RA initiative.";
RT Dev. Dyn. 225:384-391 (2002).
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC073177; AAH73177.1; -.
DR InterPro; IPR001752; kinesin_motor.
DR InterPro; IPR010994; Ruva_2_like.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1.
DR ATP-binding; Hypothetical protein; Microtubule; Motor protein.
KW SEQUENCE 631 AA; 71084 MW; 11E8C9AB86BD3F3 CRC64;
Query Match 54.9%; Score 1422; DB 2; Length 631;
Best Local Similarity 60.6%; Pred. No. 3.6e-84;
Matches 285; Conservative 78; Mismatches 99; Indels 8; Gaps 4;
QY 43 ARRVAVRLRPFDGTAGASDP-PCVVRGMDSCSLFIANRNHOETLKYQDFAFYGERSTQ 101
Db 11 ARRVAVRLRPYMBEKEEDKVTACVRLGDSLSHLSIVNWRQLETMQYQDFAFYGSASQ 70
QY 102 QDIYAGSVQPIRLHLLLEGONASVLAYGPTGAGTKHTMLGSPQPGVTPRALMDLQLTR- 160
Db 71 REIYMGSVCHILPHLLIGNASVFAYGPTGAGTKHTMLGNPDQPGVPRVRELLQWTRM 130
QY 161 -EAGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNIILPGLSKQPISSF 219
Db 131 AASAPENENWTHITMSYVEIYQEKVMDLLEPKNDLPREDKDHNIIPGVTLKINSF 190
QY 220 ADPERHFLPASNRRTVGNRLNQRSSSHAVLLVKVDORERLAPRQREGKYLIDLAGS 279
Db 191 GDFDEHFI PASQNRVASTKLNDRSSSHAVLLIKVQSQVAPRQIQLKYLIDLAGS 250
QY 280 EDNRRTGNKGLRKESGAINSLFVLGKVVVDALNOGLPRVPRDLSKLTRLLQDSLGSAH 339
Db 251 EDNRRTGNQIRLKESGAINSLFTLSKVVVDALNOGLPRIPYRDSKLTRLLQDSLGSAH 310
```

QY 340 SILIANIAPERRYLDTGVALNFAARSKVINRPFPTNESLQPHALGPVLSQKELGPPPE 399
 Db 311 SVMITNIAPEQTYFTLTALNFAAKSKQIINKPFSRETQTQVAPAMKRPREEAEATTS 370
 QY 400 AXKARGPEEEIGSPPEMAAPASAKSLPQKLSMDPAMLERLLSLDELLASQSQGA 459
 Db 371 SRORKSKSDTSESSPNWSWGKGLN----LASLDSAVVERLLKDLKILTEKGKKEA 426
 QY 460 PLLSTPKRRMVLKTVBEKDLIERLTKQKELEAKMLAQKAE-EKENH 508
 Db 427 QLLSTPKRRMALLKQWESQMEIERLKEKQKELEQKAMEAEARLEKSN 476

RESULT 12

O60845
 ID O60845 PRELIMINARY; PRT; 198 AA.
 AC O60845;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Similar to kinesin-like DNA binding protein (KID).
 GN Name=A-328A3.2;
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99425270; PubMed=10493829;
 RA Loftus B.J.; Kim U.J.; Sneddon V.P.; Kalush F.; Brandon R.,
 RA Fuhrmann J.; Mason T.; Crosby M.L.; Barnstead M.; Cronin L.,
 RA Deslattes Mays A.; Cao Y.; Xu R.X.; Kang H.L.; Mitchell S.,
 RA Eichler E.E.; Harris P.C.; Venter J.C.; Adams M.D.;
 RT "Genome duplications and other features in 12 Mb of DNA sequence from
 RT human chromosome 16p and 16q."
 RL Genomics 60:295-308(1999).
 DR EMBL; AC002301; AAC08709.1; -.
 DR HSSP; P33173; 1161.
 DR GO; GO:0005875; C:microtubule associated complex; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003774; F:Motor activity; IEA.
 DR InterPro; IPR001752; Kinesin_motor.
 DR Pfam; PF00225; Kinesin; 1.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1.
 KW ATP-binding; Microtubule; Motor protein.
 SQ SEQUENCE 198 AA; 21567 MW; 1A7FDA95B5A03B37 CRC64;

Query Match 36.5%; Score 946; DB 2; Length 198;
 Best Local Similarity 100.0%; Pred. No. 7.6e-54;
 Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 AAGSGTQQRREMAAASAAISGACRSLKIGATRRPPARVRVRLRPPFDGTAGAS 62
 Db 2 AAGSGTQQRREMAAASAAISGACRSLKIGATRRPPARVRVRLRPPFDGTAGAS 61
 QY 63 DPPCVRGMDSCSLEIANWNRHETLKYQFDAPYGERSTQDDIYAGSVQPIRLHLEGQNA 122
 Db 62 DPPCVRGMDSCSLEIANWNRHETLKYQFDAPYGERSTQDDIYAGSVQPIRLHLEGQNA 121
 QY 123 SVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQ 182
 Db 122 SVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQ 181
 QY 183 EKV 185
 Db 182 EKV 184

RESULT 13

O35232
 ID O35232 PRELIMINARY; PRT; 148 AA.
 AC O35232;

DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Kinesin motor protein Kif22 (Fragment).
 GN Name=Kif22;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97480723; PubMed=9339368;
 RA Yang Z.; Hanlon D.W.; Marszalek J.R.; Goldstein L.S.;
 RT "Identification, partial characterization, and genetic mapping of
 RT kinesin-like protein genes in mouse."
 RL Genomics 45:123-131(1997).
 DR EMBL; AF013119; AAC39968.1; -.
 DR MGD; MGI:109233; Kif22.
 DR GO; GO:0000785; C:chromatin; IEA.
 DR GO; GO:0005819; C:spindle; IEA.
 DR InterPro; IPR001752; kinesin_motor.
 DR Pfam; PF00225; Kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
 DR PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1.
 KW ATP-binding; Microtubule; Motor protein.
 FT NON_TER 1
 FT NON_TER 148
 SQ SEQUENCE 148 AA; 16596 MW; 0950C265C3B63EB7 CRC64;
 Query Match 27.2%; Score 704; DB 2; Length 148;
 Best Local Similarity 92.6%; Pred. No. 3e-38;
 Matches 137; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
 QY 133 GKTHMLGSPQGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDDPA 192
 Db 1 GKTHMLGSPQGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDDPA 60
 QY 193 SGLVIREDCRGNILIPGLSQKPSISFADFEPHFIPASNRVVGATRLNQRSRSRAVLL 252
 Db 61 SGLVIREDCRGNILIPGLTQKPIITSFDSFEQHFIPASNRVVGATRLNQRSRSRAVLL 120
 QY 253 VKVDORERLAPPRQREGKLYLIDLAGSE 280
 Db 121 VKVDORERLTPPRQREGKLYLIDLAGSE 148
 RESULT 14
 Q6L512
 ID Q6L512 PRELIMINARY; PRT; 584 AA.
 AC Q6L512;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE Putative kinesin-like DNA binding protein.
 GN Name=OJ1281_H05.2;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chow T.-Y.; Heing Y.-I.C.; Chen C.-S.; Chen H.-H.; Liu S.-M.,
 RA Chao Y.-T.; Chang S.-J.; Chen H.-C.; Chen S.-K.; Chen T.-R.,
 RA Chen Y.-L.; Cheng C.-H.; Chung C.-I.; Han S.-Y.; Hsiao S.-H.,
 RA Hsiung J.-N.; Hsu C.-H.; Huang J.-J.; Kau P.-I.; Lee M.-C.; Leu H.-L.,
 RA Li Y.-E.; Lin S.-J.; Lin Y.-C.; Wu S.-W.; Yu C.-Y.; Yu S.-W.,
 RA Wu H.-P.; Shaw J.-F.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC117265; AAT39162.1; -.
 DR InterPro; IPR010996; DNaPol_B_like.

Fri Nov 12 12:26:18 2004

```
DR InterPro: IPR000445; HHH.
DR InterPro: IPR001752; kinesin_motor.
DR Pfam: PF00633; HHH; 1.
DR Pfam: PF00225; Kinesin; 1.
DR PRINTS: PR00380; KINESINHEAVY.
DR SMART: SM00129; KISC; 1.
DR PROSITE: PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE: PS00667; KINESIN MOTOR DOMAIN2; 1.
DR ATP-binding; Microtubule; Motor protein.
KW SEQUENCE 584 AA; 64124 MW; EF829C3A15F9190F CRC64;

Query Match      24.4%; Score 633; DB 2; Length 584;
Best Local Similarity 36.1%; Pred. No. 8.5e-33;
Matches 166; Conservative 89; Mismatches 155; Indels 50; Gaps 15;

QY 31 LSKIGATRRPPPARVRVAVRLRPFDGTAGASDPPCVRGMDSCSLRIANRNHQTILK- 88
DB 1 MATAAATOSQP---VRVLRVRPHLPSEANSAEAPCVGLGSHPGGEVTVQLKDQYTSRN 57

QY 89 --YQFDAYGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 146
DB 58 ECKYLDAPFGQESRVCIEFDQESAVIPGIEGTNATVFAYGATGSGKTYTMOGTEDLPG 117

QY 147 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRNI 206
DB 118 LIPLAVSTVALCT-----GTWC-SVEISYEVYMERCYDLEPKAREIMVLDKDGNL 170

QY 207 LIPGLSQKPISSFADFERRHFLPASRNRVTGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 266
DB 171 QUKGLAWVPRSLERFHEIYSIGVQRKVAHTGLNDVSSRSHAVLSIRITT-----DV 223

QY 267 REGKLYLIDLAGEENRRTGNKGLRLKESGAINTSFLVGLKVVDALNOGLPRVYRDSKL 326
DB 224 VKGKLNLDLAGNEDNRRTCNEGIRLOESAKINQSLFALSNIYALKNKEPRIPYRESKL 283

QY 327 TRLLOSLGGSASHLILIANIAPERFYLDTVSALNFAARSKEVINRPFNTNESLOPHALGP 386
DB 284 TRILODSLGGNSHAVMIACLNPE--YQAVHTVSLAARSRHVTH-MSSASKQETPKDK 340

QY 387 VKLSQK-----ELLGPPEA-KRARGPEEEIGSPPEMAAP-----ASAKSLSPLOKL 433
DB 341 VDMAEKLRAWLESKGTKTSIQMDG-----LLSPNAIKTFLSMHKKQSGASGRVSRGKA 395

QY 434 SSMDPAMLERLL--SLDRLLAS-----QSGQAPLLSTPKR 467
DB 396 MDQDGGKIKKVLFDASAARIPAEINFHREGTQ--DIVNTTKK 433

Search completed: November 5, 2004, 18:38:41
Job time : 78.7254 secs
```

```
RESULT 15
AAT39162
ID AAT39162 PRELIMINARY; PRT; 584 AA.
AC AAT39162;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative kinesin-like DNA binding protein.
GN OUI281_H05.2.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza; Oryza sativa.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Chow T.-Y., Hsing Y.-I. C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RA "Oryza sativa BAC OUI281_H05 genomic sequence.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:11 ; Search time 51.8496 Seconds
(without alignments)
3839.562 Million cell updates/sec

Title: US-10-797-893-6

Perfect score: 1768

Sequence: 1 MGRCLSKIGATRRPPPARV.....FYLDTSALNFAARKEVIN 346

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1763	99.7	665	1 KP22 HUMAN	Q14807 homo sapien
2	1763	99.7	665	2 AAP35923	Aap35923 homo sapi
3	1572.5	88.9	660	2 Q991C7	Q991C7 mus musculu
4	1232	69.7	639	2 Q6P3R1	Q6P3R1 xenopus tro
5	1232	69.7	639	2 AAH63896	AAH63896 xenopus t
6	1202.5	68.0	651	2 Q91869	Q91869 xenopus lae
7	1202.5	68.0	651	2 AAH70549	AAH70549 xenopus l
8	1201.5	68.0	651	2 Q918K0	Q918K0 xenopus lae
9	1197.5	67.7	663	2 Q7ZYL5	Q7ZYL5 xenopus lae
10	1192.5	67.4	650	2 Q919A8	Q919A8 xenopus lae
11	1191.5	67.4	631	2 Q6GRG0	Q6GRG0 xenopus lae
12	836	47.3	198	2 Q60845	Q60845 homo sapien
13	704	39.8	148	2 Q35232	Q35232 mus musculu
14	623.5	35.3	584	2 Q6L512	Q6L512 oryza sativ
15	623.5	35.3	584	2 AAH739162	AAH739162 oryza sat
16	594.5	33.6	628	2 Q8L788	Q8L788 arabidopsis
17	594.5	33.6	664	2 Q9L288	Q9L288 arabidopsis
18	584.5	33.1	548	2 Q8SQ09	Q8SQ09 neurospora
19	578	32.7	912	2 Q7RX60	Q7RX60 neurospora
20	568.5	32.2	821	2 Q7QD56	Q7QD56 anopheles g
21	563	31.8	838	2 Q86ZB6	Q86ZB6 botrytis ci
22	559.5	31.6	784	2 Q961H5	Q961H5 drosophila
23	559	31.6	548	2 Q8N1X8	Q8N1X8 homo sapien
24	555.5	31.4	784	1 KL68 DROME	P46867 drosophila
25	552.5	31.2	898	2 Q86VS5	Q86VS5 homo sapien
26	552.5	31.2	898	2 Q8N177	Q8N177 homo sapien
27	552.5	31.2	898	2 Q7H0F3	Q7H0F3 homo sapien
28	552	31.2	1296	2 Q7RM16	Q7RM16 plasmodium
29	551.5	31.2	562	2 Q8B199	Q8B199 mus musculu
30	551.5	31.2	571	2 Q8B245	Q8B245 mus musculu
31	551.5	31.2	880	2 Q8BL11	Q8BL11 mus musculu

32 551.5 31.2 886 2 Q91WD7 Q91WD7 mus musculu
33 551.5 31.2 895 2 Q7ZUM9 Q7ZUM9 brachydanio
34 550 31.1 703 2 Q9FZ77 Q9FZ77 arabidopsis
35 547 30.9 1254 2 Q94463 Q94463 dictyostell
36 546 30.9 997 2 Q7ZXX2 Q7ZXX2 xenopus lae
37 545.5 30.9 389 2 Q7SYZ3 Q7SYZ3 xenopus lae
38 542.5 30.7 553 2 Q7PGH1 Q7PGH1 anopheles g
39 541 30.6 650 2 Q6DDE9 Q6DDE9 xenopus lae
40 540 30.5 750 2 Q7KSK2 Q7KSK2 drosophila
41 540 30.5 750 2 AAS65150 AAS65150 drosophil
42 536 30.3 784 1 KLP6_SCHPO KLP6_SCHPO
43 535.5 30.3 677 2 Q9VRK9 Q9VRK9 schizosacch
44 533.5 30.2 642 2 Q7QV55 Q7QV55 giardia lam
45 533.5 30.2 834 2 Q6PFD6 Q6PFD6 mus musculu

ALIGNMENTS

RESULT 1

KF22 HUMAN
ID KP22_HUMAN STANDARD; PRT; 665 AA.
AC Q14807; Q94814; Q9BR46;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Kinesin-like protein KIF22 (Kinesin-like DNA-binding protein)
DE (Kinesin-like protein 4).
GN Name=KIF22; Synonyms=KNSL4, KID;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96174806; PubMed=8599929;
RA Tokai N., Fujimoto-Nishiyama A., Toyoshima Y., Yonemura S.,
RA Tsukita S., Inoue J., Yamamoto T.,
RT "Kid, a novel kinesin-like DNA binding protein, is localized to
RT chromosomes and the mitotic spindle."
RL EMBO J. 15:457-467(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Lymphocytoblastoma;
MEDLINE=99009323; PubMed=9790757;
RA Song J., Murakami H., Yang Z.Q., Koga C., Adati N., Murata T.,
RA Geitinger C., Saito-Onara F., Ikeuchi T., Matsumura M., Itakura K.,
RA Kanazawa I., Sun K., Yokoyama K.K.;
RT "Human genes for KNSL4 and MAZ are located close to one another on
RT chromosome 16p11.2."
RL Genomics 52:374-377(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain, and Lung;
MEDLINE=223388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Tohiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muray D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield J.S.N., Krzywicki M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences. ";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [4]
 RP INTERACTION WITH SIAH1, AND DEGRADATION.
 RX PubMed=11146551; DOI=10.1038/97.0nc.1204002;
 RA Germani A., Bruzzoni-Giovanelli H., Fellous A., Gisselbrecht S.,
 RA Varin-Blank N., Calvo F.;
 RT "SIAH-1 interacts with alpha-tubulin and degrades the kinesin kid by
 the proteasome pathway during mitosis. ";
 RL Oncogene 19:5997-6006 (2000).
 CC -1- FUNCTION: Kinesin family that is involved in spindle formation and
 the movements of chromosomes during mitosis and meiosis. Binds to
 microtubules and to DNA.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- PTM: Ubiquitinated; mediated by SIAH1 and leading to its
 subsequent proteasomal degradation (Probable).
 CC -1- SIMILARITY: Belongs to the kinesin-like protein family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB017430; BAA33019.2; -;
 DR EMBL; AB017335; BAA33063.1; -;
 DR EMBL; AB017333; BAA33063.1; JOINED.
 DR EMBL; AB017334; BAA33063.1; JOINED.
 DR EMBL; BC004352; AAH04352.1; -;
 DR EMBL; BC028155; AAH28155.1; -;
 DR HSP; P33173; I155.
 DR Genew; HGNC:6391; KIF22.
 DR MTM; 603213; -;
 DR GO; GO:0000776; C:kinetochore; TAS.
 DR GO; GO:0005634; C:nucleus; TAS.
 DR GO; GO:0003677; F:DNA binding; TAS.
 DR GO; GO:0003777; F:microtubule motor activity; TAS.
 DR GO; GO:0007067; P:mitosis; TAS.
 DR InterPro; IPR003583; HHH 1.
 DR InterPro; IPR001752; kinesin motor.
 DR InterPro; IPR010994; Ruva_2_like.
 DR Pfam; PF00225; Kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00129; KISC; 1.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
 DR PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1.
 KW ATP-binding; Coiled coil; DNA-binding; Microtubule; Motor protein;
 Nuclear protein; Ubl conjugation.
 FT DOMAIN 40 299 Kinesin-motor.
 FT DOMAIN 465 508 Coiled coil (Potential).
 FT NP_BIND 127 134 ATP (Potential).
 FT CONFLICT 24 24 Missing (in Ref. 2).
 FT CONFLICT 122 122 S -> KV (in Ref. 2).
 FT CONFLICT 135 169 HTMLGSPQGVIPRALMDLLQLTREAGRPWA -> TH
 ACQPRATWGDPPAGHGPAAKGGCGRPANG (in Ref.
 2).
 FT CONFLICT 303 303 V -> A (in Ref. 2).
 FT CONFLICT 418 456 APASAKSLPQLKLSMDPAMRLRLSLRLASQSG
 -> SSSLCLPQPTQTEAKAATPCGAPQLGFSACLPG
 P (in Ref. 2).
 FT CONFLICT 505 513 ENHPTMLR -> RTIVPQCSG (in Ref. 2).
 FT CONFLICT 665 665 AA; 73262 MW; C6C0AC96741DD387 CRC64;
 Query Match 99.7%; Score 1763; DB 1; Length 665;
 Best Local Similarity 100.0%; Pred. No. 1.2e-128;
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 2 GRCRLSKIGATRRPPPARVAVRLRPFDVTAGASDPPCVRGMDSCSLEIANWRNHQET 61
 26 GRCRLSKIGATRRPPPARVAVRLRPFDVTAGASDPPCVRGMDSCSLEIANWRNHQET 85

62 LKQFDFAFYGERSTQODIYAGSVQPIRLHLEGONASVLAYGPTGAGKTHTMLGSPQPG 121
 86 LKQFDFAFYGERSTQODIYAGSVQPIRLHLEGONASVLAYGPTGAGKTHTMLGSPQPG 145
 122 VIPRALMDLLQLTREAGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRGN 181
 146 VIPRALMDLLQLTREAGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRGN 205
 182 LIPGLSQKPISSPADFERHFLPASRNRVTGATRLNQRSSRSRSHAVLLVKVDQRELAPFRQ 241
 206 LIPGLSQKPISSPADFERHFLPASRNRVTGATRLNQRSSRSRSHAVLLVKVDQRELAPFRQ 265
 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVVVDALNOGLPRVPYRDSKL 301
 266 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVVVDALNOGLPRVPYRDSKL 325
 302 TRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
 326 TRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 370
 RESULT 2
 AAP35923 PRELIMINARY; PRT; 665 AA.
 AC AAP35923;
 DT 02-MAR-2004 (TReMBLrel. 27, Created)
 DT 02-MAR-2004 (TReMBLrel. 27, Last sequence update)
 DE Kinesin-like 4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kalline N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
 RA Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
 RA Phelan M., Farmer A.;
 RT "Cloning of human full-length CDSs in BD Creator(TM) System Donor
 vector. ";
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 RL EMBL; BT007259; AAP35923.1; -;
 SQ SEQUENCE 665 AA; 73261 MW; C6C0AC96741DD387 CRC64;
 Query Match 99.7%; Score 1763; DB 2; Length 665;
 Best Local Similarity 100.0%; Pred. No. 1.2e-128;
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 2 GRCRLSKIGATRRPPPARVAVRLRPFDVTAGASDPPCVRGMDSCSLEIANWRNHQET 61
 26 GRCRLSKIGATRRPPPARVAVRLRPFDVTAGASDPPCVRGMDSCSLEIANWRNHQET 85
 62 LKQFDFAFYGERSTQODIYAGSVQPIRLHLEGONASVLAYGPTGAGKTHTMLGSPQPG 121
 86 LKQFDFAFYGERSTQODIYAGSVQPIRLHLEGONASVLAYGPTGAGKTHTMLGSPQPG 145
 122 VIPRALMDLLQLTREAGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRGN 181
 146 VIPRALMDLLQLTREAGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRGN 205
 182 LIPGLSQKPISSPADFERHFLPASRNRVTGATRLNQRSSRSRSHAVLLVKVDQRELAPFRQ 241
 206 LIPGLSQKPISSPADFERHFLPASRNRVTGATRLNQRSSRSRSHAVLLVKVDQRELAPFRQ 265
 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVVVDALNOGLPRVPYRDSKL 301
 266 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVVVDALNOGLPRVPYRDSKL 325
 302 TRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
 326 TRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 370

```

RESULT 3
Q99LC7 PRELIMINARY; PRT; 660 AA.
ID AC Q99LC7
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Kinesin family member 22.
GN Name=Kif22;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA Strausberg R.; TISSUE=Mammary tumor. C3;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003427; AA03427.1; -.
DR HSSP; P31173; 1155.
DR GCD; MGI:109233; Klf22.
DR GO; GO:000785; C:chromatin; IDA.
DR GO; GO:0005819; C:spindle; IDA.
DR InterPro; IPR003583; HHH_1.
DR InterPro; IPR001752; Kinesin motor.
DR InterPro; IPR010994; Ruva 2 like.
DR Pfam; PF00225; Kinesin; 1_1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00278; HHH1; 2.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Microtubule; "Motor" protein.
SQ SEQUENCE 660 AA; 73189 MW; AA8B99477BC49B3C CRC64;

Query Match 88.9%; Score 1572.5; DB 2; Length 660;
Best Local Similarity 88.4%; Pred. No. 8.5e-114;
Matches 305; Conservative 19; Mismatches 20; Indels 1; Gaps 1;

QY 2 GRCLSKIGATRPAPRVAVRLRPFDVTGAGSDPPCVRGMDSCSLIANWRHQET 61
DB 22 GRC-VSKGGLGRPPPLARVAVRLRPFDMDGETEAKELPCVRAIDSCSLVANWKYQET 80
QY 62 LKYOQDAFYGERSTQDDIYAGSVQPIILRHLEGNASVLAIGPTGAGKTHMLGSPROP 121
DB 81 LKYOQDAFYGEKSTQBEYVGSVQPIILRHLEGNASVLAIGPTGAGKTHMLGSPROP 140
QY 122 VIPRALMDLIQLTREGAEGRPWALSVTMSYLEIQEKVLDLDPASGDIVIREDCRGN 181

```

```

141 VIPRALMDLIQLAREESAEGRPMDVSVMSYLEIQEKVLDLDPASGDIVIREDCRGN 200
182 LIFGLSQKPISSPADFERHFLPASRNTVGTATRLNORSSSHAVLLVKVQDERLAPRQ 241
201 LIFGLTQKPISTFSDFEQHFLPASRNVAVGATRLNORSSSHAVLLVKVQDERLTTPRQ 260
242 REGKLYLIDLAGSEDNRRTGNKGLRLKESGAINTSLFVLGKVVVDALNQGLPRVPRYSKL 301
261 REGKLYLIDLAGSEDNRRTGNQGLRLKESGAINTSLFVLGKVVVDALNQGLPRVPRYSKL 320
302 TRLLQSLGSAHSILIANIANTAPRRFYLDVTVSALNFAARKEVIN 346
321 TRLLQSLGSAHSILIANIANTAPRRFYQDTISALNFTARSKEVIN 365

RESULT 4
Q6P3R1 PRELIMINARY; PRT; 639 AA.
ID AC Q6P3R1
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein MGC75575.
GN Name=MGC75575;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.; to the EMBL/GenBank/DBJ databases.
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC063896; AA063896.1; -.
DR InterPro; IPR001752; kinesin motor.
DR InterPro; IPR010994; Ruva 2 like.
DR Pfam; PF00225; Kinesin; 1_1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Hypothetical protein; Microtubule; Motor protein.
SQ SEQUENCE 639 AA; 71390 MW; F5664F9020A6A1AA CRC64;

Query Match 69.7%; Score 1232; DB 2; Length 639;
Best Local Similarity 68.6%; Pred. No. 2.8e-87;
Matches 236; Conservative 53; Mismatches 53; Indels 2; Gaps 1;

```


Query Match 69.7%; Score 1232; DB 2; Length 639;
 Best Local Similarity 68.6%; Pred. No. 2.8e-87;
 Matches 236; Conservative 53; Mismatches 53; Indels 2; Gaps 1;

QY 5 RLSKIGATRRPPPARVRVAVRLRPFDVGTAGASDPPCVRGMDSCSLSEIANWRNHQSTLYK 64
 DB 4 RVSILOQHKPKSSARVRVAVRLRPYMEKEDEKAPACVAGLDSQSLSEIVNWRNQLTMOY 63
 QY 65 QPDAFYGERSTOODIYAGSVQPIRLHLLLEGONASVLAGPTGAGKTHTMLGSPQPGVIP 124
 DB 64 QPDAFYGDSATQREIYMGSVCHILPHLLIGONASVFAFGPTGAGKTHTMLGNPQPGVIP 123
 QY 125 RALMDLLQLTRES--GAEGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRGNIL 182
 DB 124 RAVRDLLOMTRTAAGPENENWYITITMSYVEIYQEKVMDLLEPKNDLPPIREDKDHNIL 183
 QY 183 IGPLSQKPISSPADFERHFLPASNRVTGATRLNQRSSSRSHAVLLVKVQDORERLAPRQR 242
 DB 184 IGVQTKTINSFGDFDEHFIPIASQNRVTASTKLNDRSSSRSHAVLLIKVQSQVSPFROL 243
 QY 243 EGKLYLIDLAGSEDNRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPRYDSKLT 302
 DB 244 TGKLYLIDLAGSEDNRTGNKGLRKESGAINSSLFTLSKVVDALNOGLPRIPYRDSKLT 303
 QY 303 RLLQDSLGSAHSILIANIAPERRFYLDVTVSALNFAARSKEVIN 346
 DB 304 RLLQDSLGSTAHSMIANIAPKQYFYDTLTALNFAAKSQIIN 347

RESULT 6

QY1869 PRELIMINARY; PRT; 651 AA.
 AC QY1869; (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Kinesin (Hypothetical protein) (Chromokinesin Kkid).
 GN Name=Kkid;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20419289; PubMed=10966105;
 RA Antonio C., Ferby I., Wilhelm H., Jones M., Karsenti E., Nebreda A.R.,
 RA Vernos I.;
 RA "Xkid, a chromokinesin required for chromosome alignment on the
 RL metaphase plate.";
 RL Cell 102:425-435 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitney M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Strausberg R.;
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 KW EMBL; BC063896; AAH63896.1;
 KW Hypothetical protein.
 SQ SEQUENCE 639 AA; 71390 MW; F5664F9020A6A1AA CRC64;

QY 5 RLSKIGATRRPPPARVRVAVRLRPFDVGTAGASDPPCVRGMDSCSLSEIANWRNHQSTLYK 64
 DB 4 RVSILOQHKPKSSARVRVAVRLRPYMEKEDEKAPACVAGLDSQSLSEIVNWRNQLTMOY 63
 QY 65 QPDAFYGERSTOODIYAGSVQPIRLHLLLEGONASVLAGPTGAGKTHTMLGSPQPGVIP 124
 DB 64 QPDAFYGDSATQREIYMGSVCHILPHLLIGONASVFAFGPTGAGKTHTMLGNPQPGVIP 123
 QY 125 RALMDLLQLTRES--GAEGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRGNIL 182
 DB 124 RAVRDLLOMTRTAAGPENENWYITITMSYVEIYQEKVMDLLEPKNDLPPIREDKDHNIL 183
 QY 183 IGPLSQKPISSPADFERHFLPASNRVTGATRLNQRSSSRSHAVLLVKVQDORERLAPRQR 242
 DB 184 IGVQTKTINSFGDFDEHFIPIASQNRVTASTKLNDRSSSRSHAVLLIKVQSQVSPFROL 243
 QY 243 EGKLYLIDLAGSEDNRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPRYDSKLT 302
 DB 244 TGKLYLIDLAGSEDNRTGNKGLRKESGAINSSLFTLSKVVDALNOGLPRIPYRDSKLT 303
 QY 303 RLLQDSLGSAHSILIANIAPERRFYLDVTVSALNFAARSKEVIN 346
 DB 304 RLLQDSLGSTAHSMIANIAPKQYFYDTLTALNFAAKSQIIN 347

RESULT 5

AAH63896 PRELIMINARY; PRT; 639 AA.
 AC AAH63896;
 DT 25-MAR-2004 (TrEMBLrel. 27, Created)
 DT 25-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 25-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein MGC75575.
 GN MGC75575.
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus; Silurana.
 OC NCBI_TaxID=8364;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitney M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Strausberg R.;
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 KW EMBL; BC063896; AAH63896.1;
 KW Hypothetical protein.
 SQ SEQUENCE 639 AA; 71390 MW; F5664F9020A6A1AA CRC64;

QY 5 RLSKIGATRRPPARVAVRLRPFVDGTAGA-SDDPCVRGMDSCSLEIANWRNHQETLK 63
 DB 17 RVSMLDQHKSSCARVAVRLRPFVDGTAGA-SDDPCVRGMDSCSLEIANWRNHQETLK 76
 QY 64 YQDFAFYGBRSTQDDIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPGVI 123
 DB 77 YQDFAFYGBRSTQDDIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPGVI 136
 QY 124 PRALMDLLOLTREEGA--EGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRNI 181
 DB 137 PRAVRDLLOLTREEGA--EGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRNI 196
 QY 182 LIPGLSOKPISSPADFERHFLPASNRRTVGATRLNORSSRSRSHAVLLVKVDQERLAPFRQ 241
 DB 197 LIPGVTKMINSPADFERHFLPASNRRTVGATRLNORSSRSRSHAVLLVKVDQERLAPFRQ 256
 QY 242 REGKLYLIDLAGSEDRNRGTGKRLKESGAINTSFLVKVVDALNOGLPRVPRYRDSKL 301
 DB 257 LTGKLYLIDLAGSEDRNRGTGKRLKESGAINTSFLVKVVDALNOGLPRVPRYRDSKL 316
 QY 302 TRLLQDSLGSAHSILITANIAPERFRFYDLTVSALNFAARKEVIN 346
 DB 317 TRLLQDSLGSAHSILITANIAPERFRFYDLTVSALNFAARKEVIN 361

RESULT 8
 QY18K0 PRELIMINARY; PRT; 651 AA.
 AC QY18K0;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Chromokinesin Xkid
 OS Xenopus laevis (African clawed frog)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20419286; PubMed=10966104;
 RA Funabiki H., Murray A.W.;
 RT "The Xenopus chromokinesin xkid is essential for metaphase chromosome alignment and must be degraded to allow anaphase chromosome movement";
 RT Cell 102:411-424 (2000).
 RL EMBL; AF267850; AAF82564.1; --
 DR HSSP; P33173; 115S.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0005875; C:microtubule associated complex; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003774; F:motor activity; IEA.
 DR InterPro; IPR010996; DNAPOL_B_N-like.
 DR InterPro; IPR003583; HHH 1.
 DR Pfam; PF00225; Kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00278; HhH1; 2.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
 DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
 KW ATP-binding; Microtubule; Motor protein.
 FT NON_TER 1
 SQ SEQUENCE 651 AA; 73091 MW; C4EF79801E603C13 CRC64;

Query Match 68.0%; Score 1201.5; DB 2; Length 651;
 Best Local Similarity 67.8%; Pred. No. 6.8e-85;
 Matches 234; Conservative 52; Mismatches 56; Indels 3; Gaps 2;
 QY 5 RLSKIGATRRPPARVAVRLRPFVDGTAGA-SDDPCVRGMDSCSLEIANWRNHQETLK 63
 DB 17 RVSMLDQHKSSCARVAVRLRPFVDGTAGA-SDDPCVRGMDSCSLEIANWRNHQETLK 76

QY 64 YQDFAFYGBRSTQDDIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPGVI 123
 DB 77 YQDFAFYGBRSTQDDIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPGVI 136
 QY 124 PRALMDLLOLTREEGA--EGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRNI 181
 DB 137 PRAVRDLLOLTREEGA--EGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRNI 196
 QY 182 LIPGLSOKPISSPADFERHFLPASNRRTVGATRLNORSSRSRSHAVLLVKVDQERLAPFRQ 241
 DB 197 LIPGVTKMINSPADFERHFLPASNRRTVGATRLNORSSRSRSHAVLLVKVDQERLAPFRQ 256
 QY 242 REGKLYLIDLAGSEDRNRGTGKRLKESGAINTSFLVKVVDALNOGLPRVPRYRDSKL 301
 DB 257 LTGKLYLIDLAGSEDRNRGTGKRLKESGAINTSFLVKVVDALNOGLPRVPRYRDSKL 316
 QY 302 TRLLQDSLGSAHSILITANIAPERFRFYDLTVSALNFAARKEVIN 346
 DB 317 TRLLQDSLGSAHSILITANIAPERFRFYDLTVSALNFAARKEVIN 361

RESULT 9

QY7ZYL5 PRELIMINARY; PRT; 663 AA.
 AC QY7ZYL5;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Similar to kinesin family member 22 (Fragment).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC043733; AAH43733.1; --
 DR HSSP; P33173; 115S.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0005875; C:microtubule associated complex; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003774; F:motor activity; IEA.
 DR InterPro; IPR010996; DNAPOL_B_N-like.
 DR InterPro; IPR003583; HHH 1.
 DR InterPro; IPR001752; kinesin_motor.
 DR Pfam; PF00225; Kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00278; HhH1; 2.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
 DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
 KW ATP-binding; Microtubule; Motor protein.
 FT NON_TER 1
 SQ SEQUENCE 663 AA; 74538 MW; 3B98B5F3C12C9C22 CRC64;

Query Match 67.7%; Score 1197.5; DB 2; Length 663;
 Best Local Similarity 67.5%; Pred. No. 1.4e-84;
 Matches 233; Conservative 52; Mismatches 57; Indels 3; Gaps 2;
 QY 5 RLSKIGATRRPPARVAVRLRPFVDGTAGA-SDDPCVRGMDSCSLEIANWRNHQETLK 63
 DB 30 RVSMLDQHKSSCARVAVRLRPFVDGTAGA-SDDPCVRGMDSCSLEIANWRNHQETLK 89

QY 64 YQDFAFYGBRSTQDDIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPGVI 123
 DB 90 YQDFAFYGBRSTQDDIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPGVI 149
 QY 124 PRALMDLLOLTREEGA--EGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRNI 181

```

Db 150 PRVRELLQWTRMAASAPENENWTTITMSYVEIYQEKVMDLLEPKNDLP IREDKDHNI 209
Qy 182 LIPGLSOKPISSFADFERHFLPASRNRRTVGATRLNORSRSHAVLLVKVDQERLAPRQ 241
Db 210 LIPGVTLTKTNSFGDFEHPFIPASQNRFTVASTKLNDRSSRSHAVLLIKVKQSQVAPRQ 269
Qy 242 REGKLYLIDLAGSEBDRRTGNKGLRKESGAINSTSLFVLGKVVDALNOGLPRVPYRDSKL 301
Db 270 LIGKLYLIDLAGSEBDRRTGNQIRLKESGAINSTSLFSLSKVDALNOGLPRIPYRDSKL 329
Qy 302 TRLLQDSLGGSAAHILLANIAPERRFYLDTVSALNFAARKEVIN 346
Db 330 TRLLQDSLGGSAAHVMITNIAPEQTYFYFDLTALNFAAKSKQIIN 374

RESULT 10
Q919A8 PRELIMINARY; PRT; 650 AA.
ID AC Q919A8
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Kinesin.
GN Name=kid;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20419289; PubMed=10966105;
RA Antonio C., Ferby I., Wilhelm H., Jones M., Karsenti E., Nebreda A.R.,
RA Vernos I.;
RT "Xkid, a chromokinesin required for chromosome alignment on the
RT metaphase plate.";
RL Cell 102:425-435 (2000).
DR EMBL; AJ249841; CAB71799.1; -.
DR HSP; P33173; 1155.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005875; C:microtubule associated complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR InterPro; IPR010996; DNAPOL_B_N-like.
DR InterPro; IPR003583; HHH_1.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00278; HhH1; 2.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.
DR ATP-binding; Microtubule; Motor protein.
SQ SEQUENCE 650 AA; 73027 MW; EA8415BC4B9B0F4A CRC64;

Query Match 67.48; Score 1192.5; DB 2; Length 650;
Best Local Similarity 67.58; Pred.No.3.4e-84;
Matches 233; Conservative 51; Mismatches 58; Indels 3; Gaps 2;

Qy 5 RLSKIGATRRPPPARVAVRLRPFDVGTAGADP-PCVRGMDSCSLEIANWRNHQDTLK 63
Db 17 RASMLDQHKSSCARVAVRLRPYMEKEEDKVPTACVGLDSSHLSLEIYVNWENQLETMQ 76
Qy 64 YQFDAPFGRSTQDDIYAGSVQPIRLHLEGONASVLAYGPTGAGKTHTMLGSPQPGVI 123
Db 77 YQFDAPFYGDSQREIYMGVSVCHILPHLLIQONASVFAYGPTGAGKTHTMLGNPDQPGVI 136
Qy 124 PRALMDLQLQTR--EEGAERPRWALSVTMSYLEIYQEKVMDLLEPKNDLP IREDKDHNI 191
Db 137 PRVRELLQWTRMAASAPENENWTTITMSYVEIYQEKVMDLLEPKNDLP IREDKDHNI 196

```

```

Qy 182 LIPGLSOKPISSFADFERHFLPASRNRRTVGATRLNORSRSHAVLLVKVDQERLAPRQ 241
Db 197 LIPGVTLTKTNSFGDFEHPFIPASQNRFTVASTKLNDRSSRSHAVLLIKVKQSQVAPRQ 256
Qy 242 REGKLYLIDLAGSEBDRRTGNKGLRKESGAINSTSLFVLGKVVDALNOGLPRVPYRDSKL 301
Db 257 LIGKLYLIDLAGSEBDRRTGNQIRLKESGAINSTSLFSLSKVDALNOGLPRIPYRDSKL 316
Qy 302 TRLLQDSLGGSAAHILLANIAPERRFYLDTVSALNFAARKEVIN 346
Db 317 TRLLQDSLGGSAAHVMITNIAPEQTYFYFDLTALNFAAKSKQIIN 361

RESULT 11
Q6GPG0 PRELIMINARY; PRT; 631 AA.
ID AC Q6GPG0
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield V.S.,
RA Kzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC073177; AAH73177.1; -.
DR InterPro; IPR01752; kinesin_motor.
DR InterPro; IPR010994; RuvA_2_like.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.
DR ATP-binding; Hypothetical protein; Microtubule; Motor protein.
KW ATP-binding; Hypothetical protein; Microtubule; Motor protein.

```

SQ SEQUENCE 631 AA; 71084 MW; 11E8C9AB86EBD3F3 CRC64;
Query Match 67.4%; Score 1191.5; DB 2; Length 631;
Best Local Similarity 69.6%; Pred. No. 3.9e-84;
Matches 231; Conservative 48; Mismatches 50; Indels 3; Gaps 2;
QY 18 ARRVAVRLRPVVDGTAGASDP-PCVRGMDSCSLSIANRNHOETLKQYQDAFYGERSTQ 76
DB 11 ARRVAVRLRPVVEEKEDKVPCTACVGRGLDSDLSLEIVNWRNQLTMTQYQDAFYGDSASQ 70
QY 77 QDIYAGSVQPIRLHLEQNASVLAGPTGAGKTHMLGSPQGPVPIRALMDLLQLTR- 135
DB 71 REIYMGSVCHILPHLLIGQNASVFAYGPTGAGKTHMLGNPDQGPVIFRAVRELLQWTRM 130
QY 136 -EAGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGNILIPGLSQPISF 194
DB 131 AASAPENWNTYITMSYVEIYQEKVMDLLEPKNDLPREDKDHNLIPGVTLKTNSF 190
QY 195 ADPERHFLPASNRVTGATRLNORSRSHAVLLVKVDQERLAPFQREGKLYLIDLGS 254
DB 191 GDFDEHIFPASQNRVTASTKLANDRSSRSHAVLLIKVQSQVAPFQRLIGKLYLIDLGS 250
QY 255 EDNRRTGNKLRKESGANTSLFVLGKVDALNOGLPRVPYRDSKLTRELQDSLGSAH 314
DB 251 EDNRRTGNQIRUKESGAINSSFTLSKVVDALNQGLPRIPYRDSKLTRELQDSLGSAH 310
QY 315 SILIANIAPERFYDLTVALNPAARSKVIN 346
DB 311 SVMITHIAPEQTYFTLTALNPAARSKVIN 342
RESULT 12
O60845 PRELIMINARY; PRT; 198 AA.
AC O60845
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to kinesin-like DNA binding protein (KID).
GN Name=A-328A3.2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=9445270; PubMed=10493829;
RA Loftus B.J., Kim U.J., Sneddon V.P., Kalush F., Brandon R.,
RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.L., Mitchell S.,
RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
RT "Genome duplications and other features in 12 Mb of DNA sequence from
human chromosome 16p and 16q."
RL Genomics 60:295-308(1999).
DR EMBL; AC002301; AAC08709.1; -.
DR HSRP; P33173.1; I61.
DR GO; GO:0005875; C:microtubule associated complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR InterPro; IPR001752; Kinesin_motor.
DR Pfam; PF00225; kinesin_1.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.
SQ SEQUENCE 198 AA; 21567 MW; 1A7F0A95B5A03B37 CRC64;

Query Match 47.3%; Score 836; DB 2; Length 198;
Best Local Similarity 100.0%; Pred. No. 4.2e-57;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GRCRLSKIGATRRPPPARVAVRLRPVVDGTAGASDPCCVRGMDSCSLSIANRNHOET 61
DB 26 GRCRLSKIGATRRPPPARVAVRLRPVVDGTAGASDPCCVRGMDSCSLSIANRNHOET 85

QY 62 LKQYQDAFYGERSTQODIYAGSVQPIRLHLEQNASVLAGPTGAGKTHMLGSPQGP 121
DB 86 LKQYQDAFYGERSTQODIYAGSVQPIRLHLEQNASVLAGPTGAGKTHMLGSPQGP 145
QY 122 VIPRALMDLLQLTRREGAGRPWALSVTMSYLEIYQEKV 160
DB 146 VIPRALMDLLQLTRREGAGRPWALSVTMSYLEIYQEKV 184
RESULT 13
O35232 PRELIMINARY; PRT; 148 AA.
AC O35232
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Kinesin motor protein KIF22 (Fragment).
GN Name=Kif22;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=97480723; PubMed=9339368;
RA Yang Z., Hanlon D.W., Marszalek J.R., Goldstein L.S.;
RT "Identification, partial characterization, and genetic mapping of
kinesin-like protein genes in mouse."
RL Genomics 45:123-131(1997).
DR EMBL; AF013119; AAC39968.1; -.
DR MGD; MGI:109233; Kif22.
DR GO; GO:0000785; C:chromatin; IEA.
DR GO; GO:0005819; C:spindle; IEA.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin_1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.
FT NON_TER 1
FT NON_TER 148
FT NON_TER 148
SQ SEQUENCE 148 AA; 16596 MW; 0950C265C3B63EB7 CRC64;
Query Match 39.8%; Score 704; DB 2; Length 148;
Best Local Similarity 92.6%; Pred. No. 5.5e-47;
Matches 137; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
QY 108 GKXTHMLGSPQGPVPIRALMDLLQLTRREGAGRPWALSVTMSYLEIYQEKVLDLDP 167
DB 1 GKXTHMLGSPQGPVPIRALMDLLQLTRREGAGRPWALSVTMSYLEIYQEKVLDLDP 60
QY 168 SGDLVIREDCRGNILIPGLSQPISFADFERHFLPASNRVTGATRLNORSRSHAVLL 227
DB 61 SGDLVIREDCRGNILIPGLTQKPTITFSDFDFQFLPASNRVAVGATRLNORSRSHAVLL 120
QY 228 VKVDQERLAPFQREGKLYLIDLGS 255
DB 121 VKVDQERLTPFQREGKLYLIDLGS 148
RESULT 14
Q6L512 PRELIMINARY; PRT; 584 AA.
AC Q6L512
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative kinesin-like DNA binding protein.
GN Name=CJ1281.H05.2;
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
 RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
 RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
 RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C.,
 RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
 RA Wu H.-P., Shaw J.-F.,
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC117265; AAT39162.1; -
 DR InterPro; IPR010996; DNaPol_B_N-like.
 DR InterPro; IPR000445; HH.
 DR InterPro; IPR001752; kinesin_motor.
 DR Pfam; PF00633; HHH; 1.
 DR Pfam; PF00225; Kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
 DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
 KW ATP-binding; Microtubule; Motor protein.
 SQ SEQUENCE 584 AA; 64124 MW; EF829C3A15F9190F CRC64;
 Query Match 35.3%; Score 623.5; DB 2; Length 584;
 Best Local Similarity 40.9%; Pred. No. 6.4e-40;
 Matches 141; Conservative 68; Mismatches 113; Indels 23; Gaps 7;
 QY 6 LSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPCCVGRMDS-CSLEIANWRNHQETLK- 63
 Db 1 MATAAATQSQP---VRVVLVRPHLPSEANSAEAPCVGLLGSHPGGEVTVQLKDQYTSRN 57
 QY 64 --YQFDAPYGRSTQDDIYAGSVQPIILRHLEGGONASVLAAYGPTGAGKTHMLGSPQPG 121
 Db 58 ECKYKLDAPFGQESRVCEIFDQEVSAVIGPIEGTNATVFAYGATGSGKTYTMOGTEDLPG 117
 QY 122 VIPALMDLLQLTREGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDIVREDRCGNI 181
 Db 118 LIPLAVSTVLALCT-----GTWC-SVEISYIYEMERCYDLLEPKAREIMVLDKQGNL 170
 QY 182 LIPGLSKPISSFADFERRHFLPASRNRTVGATRLNQSRSSRHAVLLVKVDQRLAPFRQ 241
 Db 171 QLKGLAWPVRSLEEFHEIYSIGVQRKVAHTGLNDVSSRSHAVLSIRITT-----DV 223
 QY 242 REGKLYLIDLAGSDNRRTGNKGLKESGAINTSLFVLGKVVDALNQGLPRVPYRDSKL 301
 Db 224 VKGKLNILDLAGNEDNRTCNEGIRLQESAKINQSLFALSNNVISALNKKKEPRIPYRESKL 283
 QY 302 TRILQDSLGGSAHSILIANIAPERRFYLDITVSALNFAARKEVIN 346
 Db 284 TRILODSLGGNSHAVMIACLNPVE--YQEAHVTVSLAARSHVTN 326
 RESULT 15
 AAT39162
 ID AAT39162 PRELIMINARY; PRT; 584 AA.
 AC AAT39162;
 DT 01-JUN-2004 (TrEMBLrel. 27, Created)
 DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
 DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
 DE Putative kinesin-like DNA binding protein.
 GN OJ1281.H05.2.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzae; Oryza; Oryza sativa.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
 RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,

RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
 RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C.,
 RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
 RA Wu H.-P., Shaw J.-F.,
 RT "Oryza sativa BAC OJ1281.H05 genomic sequence."
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC117265; AAT39162.1; -
 SQ SEQUENCE 584 AA; 64124 MW; EF829C3A15F9190F CRC64;
 Query Match 35.3%; Score 623.5; DB 2; Length 584;
 Best Local Similarity 40.9%; Pred. No. 6.4e-40;
 Matches 141; Conservative 68; Mismatches 113; Indels 23; Gaps 7;
 QY 6 LSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPCCVGRMDS-CSLEIANWRNHQETLK- 63
 Db 1 MATAAATQSQP---VRVVLVRPHLPSEANSAEAPCVGLLGSHPGGEVTVQLKDQYTSRN 57
 QY 64 --YQFDAPYGRSTQDDIYAGSVQPIILRHLEGGONASVLAAYGPTGAGKTHMLGSPQPG 121
 Db 58 ECKYKLDAPFGQESRVCEIFDQEVSAVIGPIEGTNATVFAYGATGSGKTYTMOGTEDLPG 117
 QY 122 VIPALMDLLQLTREGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDIVREDRCGNI 181
 Db 118 LIPLAVSTVLALCT-----GTWC-SVEISYIYEMERCYDLLEPKAREIMVLDKQGNL 170
 QY 182 LIPGLSKPISSFADFERRHFLPASRNRTVGATRLNQSRSSRHAVLLVKVDQRLAPFRQ 241
 Db 171 QLKGLAWPVRSLEEFHEIYSIGVQRKVAHTGLNDVSSRSHAVLSIRITT-----DV 223
 QY 242 REGKLYLIDLAGSDNRRTGNKGLKESGAINTSLFVLGKVVDALNQGLPRVPYRDSKL 301
 Db 224 VKGKLNILDLAGNEDNRTCNEGIRLQESAKINQSLFALSNNVISALNKKKEPRIPYRESKL 283
 QY 302 TRILQDSLGGSAHSILIANIAPERRFYLDITVSALNFAARKEVIN 346
 Db 284 TRILODSLGGNSHAVMIACLNPVE--YQEAHVTVSLAARSHVTN 326
 Search completed: November 5, 2004, 18:38:42
 Job time : 52.8496 secs

This page blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:11 ; Search time 72.979 Seconds
(without alignments)
3839.562 Million cell updates/sec

Title: US-10-797-893-8

Perfect score: 2472

Sequence: 1 MGRCLSKIGATRRPPARV.....LEAKMLAQAEKENHCPTM 487

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2467	99.8	665	1 KF22 HUMAN	Q14807 homo sapien
2	2467	99.8	665	2 AAP35923	Aap35923 homo sapi
3	2129.5	86.1	660	2 Q91C7	Q91C7 mus musculus
4	1463	59.2	639	2 Q6P31	Q6P31 xenopus tro
5	1463	59.2	639	2 AAH63896	AAH63896 xenopus t
6	1438	58.2	651	2 Q91869	Q91869 xenopus lae
7	1438	58.2	651	2 AAH70549	AAH70549 xenopus l
8	1437	58.1	651	2 Q918K0	Q918K0 xenopus lae
9	1428	57.8	663	2 Q7ZYL5	Q7ZYL5 xenopus lae
10	1423	57.6	650	2 Q919A8	Q919A8 xenopus lae
11	1422	57.5	631	2 Q6GP00	Q6GP00 xenopus lae
12	836	33.8	198	2 Q60845	Q60845 homo sapien
13	704	28.5	148	2 Q35232	Q35232 mus musculus
14	633	25.6	584	2 Q6L512	Q6L512 oryza sativ
15	633	25.6	584	2 AAT39162	AAT39162 oryza sat
16	616	24.9	784	2 Q961H5	Q961H5 drosophila
17	614	24.8	628	2 Q8L788	Q8L788 arabidopsis
18	614	24.8	664	2 Q9L288	Q9L288 arabidopsis
19	612	24.8	784	1 KL68_DROME	P46867 drosophila
20	611.5	24.7	838	2 Q86ZB6	Q86ZB6 botrytis ci
21	608.5	24.6	548	2 Q8SQ09	Q8SQ09 encephalito
22	604	24.4	912	2 Q7RX60	Q7RX60 neuropept
23	603.5	24.4	821	2 Q7QDS6	Q7QDS6 anopheles g
24	597.5	24.2	703	2 Q9FZ77	Q9FZ77 arabidopsis
25	593	24.0	677	2 Q9VRK9	Q9VRK9 drosophila
26	594.5	23.6	642	2 Q7QV55	Q7QV55 giardia lam
27	590.5	23.5	1394	2 Q7M624	Q7M624 mus musculus
28	579.5	23.4	1226	1 KF4A_XENIA	Q1784 xenopus lae
29	579.5	23.4	1226	2 Q6IRM2	Q6IRM2 xenopus lae
30	579.5	23.4	1226	2 AAH70854	AAH70854 xenopus l
31	579	23.4	1463	2 Q9GY20	Q9GY20 strongyloce

32 577 23.3 571 2 Q8BZ45 Q8BZ45 mus musculus
33 577 23.3 880 2 Q8BL11 Q8BL11 mus musculus
34 577 23.3 886 2 Q91WD7 Q91WD7 mus musculus
35 575.5 23.3 1231 2 Q6DINS Q6DINS xenopus tro
36 574.5 23.2 787 2 Q6PKB2 Q6PKB2 homo sapien
37 574.5 23.2 787 2 AAH03664 AAH03664 homo sapi
38 574.5 23.2 1127 2 Q86TN3 Q86TN3 homo sapien
39 574.5 23.2 1232 1 KF4A_HUMAN Q85239 homo sapien
40 574 23.2 562 2 Q8BY99 Q8BY99 mus musculus
41 572 23.1 895 2 Q7ZUW9 Q7ZUW9 brachydanio
42 570 23.1 699 2 Q98T11 Q98T11 xenopus lae
43 569.5 23.0 915 2 Q6GR48 Q6GR48 xenopus lae
44 568.5 23.0 671 2 Q86XX7 Q86XX7 homo sapien
45 568 23.0 898 2 Q86VS5 Q86VS5 homo sapien

ALIGNMENTS

RESULT 1
KF22 HUMAN
ID KF22 HUMAN STANDARD; PRT; 665 AA.
AC Q14807; Q94814; Q9BT46;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DS Kinesin-like protein KIF22 (Kinesin-like DNA-binding protein)
DE (Kinesin-like protein 4).
GN Name=KIF22; Synonyms=KNSL4, KID;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96174806; PubMed=8599929;
TX Tokai N., Fujimoto-Nishiyama A., Toyoshima Y., Yonemura S.,
RA Taukita S., Inoue J., Yamamoto T.;
RT "Kid, a novel kinesin-like DNA binding protein, is localized to
RT chromosomes and the mitotic spindle.";
RL EMBO J. 15:457-467(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Lymphocytoblastoma;
MEDLINE=99009323; PubMed=9790757;
RA Song J., Murakami H., Yang Z.Q., Koga C., Adati N., Murata T.,
RA Geltinger C., Saito-Ohara F., Ikeuchi T., Matsumura M., Itakura K.,
RA Kanazawa I., Sun K., Yokoyama K.K.;
RT "Human genes for KNSL4 and MAZ are located close to one another on
RL chromosome 16p11.2.";
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain, and Lung;
MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uadin T.B., Tohiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

QY 182 LIPGLSQPISSPADFERHFLPASGRNRTVGATRLNQRSSRSHAVLLVKVDQERLAPRQ 241
 Db 206 LIPGLSQPISSPADFERHFLPASGRNRTVGATRLNQRSSRSHAVLLVKVDQERLAPRQ 265
 QY 242 REGKYLIDLAGSDNRRTGNKGLRKGSGAINTSLFVLGKVDALNOGLPRVYRDSKL 301
 Db 266 REGKYLIDLAGSDNRRTGNKGLRKGSGAINTSLFVLGKVDALNOGLPRVYRDSKL 325
 QY 302 TRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARSKVEINRPTNESLOPHALGP 361
 Db 326 TRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARSKVEINRPTNESLOPHALGP 385
 QY 362 VKLSQKELLGPPPEAKRARGPPEBEIGSPPEPMAAPASASQKLSPLQKSSMDPAMLERLLS 421
 Db 386 VKLSQKELLGPPPEAKRARGPPEBEIGSPPEPMAAPASASQKLSPLQKSSMDPAMLERLLS 445
 QY 422 LDRLLASQSGCAPLLSTPKRERVMKMTVEEKDLEIRLTKOKELAKMLAOKAKE 481
 Db 446 LDRLLASQSGCAPLLSTPKRERVMKMTVEEKDLEIRLTKOKELAKMLAOKAKE 505
 QY 482 NHCPTM 487
 Db 506 NHCPTM 511

RESULT 3
 Q99LC7
 ID Q99LC7 PRELIMINARY; PRT; 660 AA.
 AC Q99LC7
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Kinesin family member 22.
 GN Name=Kif22;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RX MEDLINE=223825; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Hax S.I., Wang J., Heieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Frange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Kzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RC SEQUENCE FROM N.A.
 RP STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RX Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RA EMBL; BC003427; AA03427.1; -.
 DR HSP; P33173; 1155.
 DR MGD; MGI:109233; Kif22.
 DR GO; GO:0000785; C:chromatin; IDA.
 DR GO; GO:0005819; C:spindle; IDA.
 DR InterPro; IPR003583; HHH_1.

DR InterPro; IPR001752; kinesin_motor.
 DR InterPro; IPR010994; RUVA_2_like.
 DR Pfam; PF00225; Kinesin; 1_...
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00278; HHH1; 2.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
 DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
 KW ATP-binding; Microtubule; Motor protein.
 SQ SEQUENCE 660 AA; 73189 MW; AA8B99477BC49B3C CRC64;

Query Match 86.1%; Score 2129.5; DB 2; Length 660;
 Best Local Similarity 86.3%; Pred. No. 5.5e-131;
 Matches 422; Conservative 29; Mismatches 33; Indels 5; Gaps 3;

QY 2 GRCLSLKIGATRRPPPARVAVRLRPVVDGTAGACPPCVRGWDSCLSETANRNHOET 61
 Db 22 GRC-VSGKGLRGRLPRLARVAVRLRPFMDGETEAKELPCVRAIDSCSLEVANWKYQET 80
 QY 62 LKQFDIFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLAIGPTGAGKTHMLGSPQPG 121
 Db 81 LKQFDIFYGEKSTQQEVVGVQPIRLHLLLEGQNASVLAIGPTGAGKTHMLGSPQPG 140
 QY 122 VIPRALMDLLQTLREAGRGPRWALSVTMSYLEIYQEKVLDLDPASGDVLRDRCRNI 181
 Db 141 VIPRALMDLLQTLAREESAEGRPMDVSVAMSYLEYQEKVLDLDPASGDVLRDRCRNI 200
 QY 182 LIPCLSKPISSPADFERHFLPASGRNRTVGATRLNQRSSRSHAVLLVKVDQERLAPRQ 241
 Db 201 LIPGLTQKPIITSPFDFEQHFLPASRNRVAVGATRLNQRSSRSHAVLLVKVDQERLTPPRQ 260
 QY 242 REGKYLIDLAGSDNRRTGNKGLRKGSGAINTSLFVLGKVDALNOGLPRVYRDSKL 301
 Db 261 REGKYLIDLAGSDNRRTGNKGLRKGSGAINTSLFVLGKVDALNOGLPRVYRDSKL 320
 QY 302 TRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARSKVEINRPTNESLOPHALGP 361
 Db 321 TRLLQDSLGGSAHSILIANIAPERRFYQDTISALNPTARSKEVINRPTNESLOPHALAP 380
 QY 362 VKLSQKELLGPPPEAKRARGPPEBEIGSPPEPMAAPASASQKLSPLQKSSMDPAMLERLLS 421
 Db 381 VKLSQKELLGPPPEAKRARGPPEBEIGSPPEPMAAPASASQKLSPLQKSSMDPAMLERLLS 440
 QY 422 LDRLLASQSGCAPLLSTPKRERVMKMTVEEKDLEIRLTKOKELAKMLAOKAKE 478
 Db 441 MERLLSQSGQGTPLNTPKRERVMKMTVEEKDLEIRLTKOKELAKMLAOKAKE 500
 QY 479 EKENHCPMT 487
 Db 501 EKEN-TPTI 508

RESULT 4
 Q6P3R1
 ID Q6P3R1 PRELIMINARY; PRT; 639 AA.
 AC Q6P3R1
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 03-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE Hypothetical protein MGC75575.
 GN Name=MGC75575;
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8364;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP TISSUE=Embryo;
 RX MEDLINE=223825; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.; to the EMBL/GenBank/DBJ databases.
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC063896; AAH63896.1; -
DR InterPro; IPR01752; Kinesin motor.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS04411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Hypothetical protein; Microtubule; Motor protein.
SQ SEQUENCE 639 AA; 71390 MW; F5664F9020A6A1AA CRC64;

Query Match 59.2%; Score 1463; DB 2; Length 639;
Best Local Similarity 60.3%; Pred. No. 2.4e-87;
Matches 292; Conservative 79; Mismatches 103; Indels 10; Gaps 4;

QY 5 RLSKIGATRRPPARVAVRLRPFVDTAGASDPPCVRGMDSCSLEIANWRHQTLY 64
DB 4 RVSLDQHKPPSARVAVRLRPFVDTAGASDPPCVRGMDSCSLEIANWRHQTLY 63

QY 65 QFDAYGERSTQDIYAGSVQPIRLHLEGONASVLAAYGTGAGKTHMLGSPQGVIP 124
DB 64 QFDAYGDSATQREIYMGSVCHILPHLLIGONASVFAVGTGAGKTHMLGNPSQGVIP 123

QY 125 RALMDLLQLTREE--GAEGRPWALSVMYSLEYQEKVLDLDPASGDLVIREDCRGNIL 182
DB 124 RAVRDLLQMTTAAGPENENWTITMSYVEIYQEKVMDLLEPKNKDLPREDKDHNIL 183

QY 183 IGPLSQKPISSPADFERHFLPASNRVTGATRLNQRSSRSHAVLLVKVQDRELAPROR 242
DB 184 IPGVTQKTINSFGDFDEHFIPASQNTVASTKLNDRSSRSHAVLLIKVQSQVSPRQL 243

QY 243 EGKLYLIDLAGESENRRTGNKGLRKESGAINTSFLVLGKVVDALNQLPRVYRDSKLT 302
DB 244 TGKLYLIDLAGESENRRTGNKGLRKESGAINTSFLVLGKVVDALNQLPRVYRDSKLT 303

QY 303 RLLQDSLGSASHILIANIAPERFDYTSALNFAARKEVINRPTNSLQPHALGPV 362
DB 304 RLLQDSLGSASHILIANIAPERFDYTSALNFAARKEVINRPTNSLQPHALGPV 363

QY 363 KLSQKELLGPENKARPEEEIGSPPEMAAPASAKLSPLQK--LSSMDPAMLERLL 420
DB 364 MKRPRE-----EAETAAGSRQKKSKTSTESSNTMDAASKRLNLAALDPVAVRELL 418

QY 421 SLDRLLASQGSOCAPLLTPKRRMVLMTVEEKDLERLTKTKQLEAKQAKAE-E 479
DB 419 KLDKILTEGKKEAQLLSTPKRRMVLMTVEEKDLERLTKTKQLEAKAEAEARLE 478

QY 480 KENH 483
DB 479 KSTN 482

RESULT 5
AAH63896 PRELIMINARY; PRT; 639 AA.
AC AAH63896;
DT 25-MAR-2004 (T-REMBLrel. 27, Created)
DT 25-MAR-2004 (T-REMBLrel. 27, Last sequence update)
DT 25-MAR-2004 (T-REMBLrel. 27, Last annotation update)
DE Hypothetical protein MGC75575.
GN MGC75575.
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodidae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
EX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Shevchenko Y., Bouffard G.G.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Skalska U., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.; to the EMBL/GenBank/DBJ databases.
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC063896; AAH63896.1; -
KW Hypothetical protein.
SQ SEQUENCE 639 AA; 71390 MW; F5664F9020A6A1AA CRC64;

Query Match 59.2%; Score 1463; DB 2; Length 639;
Best Local Similarity 60.3%; Pred. No. 2.4e-87;
Matches 292; Conservative 79; Mismatches 103; Indels 10; Gaps 4;

QY 5 RLSKIGATRRPPARVAVRLRPFVDTAGASDPPCVRGMDSCSLEIANWRHQTLY 64
DB 4 RVSLDQHKPPSARVAVRLRPFVDTAGASDPPCVRGMDSCSLEIANWRHQTLY 63

QY 65 QFDAYGERSTQDIYAGSVQPIRLHLEGONASVLAAYGTGAGKTHMLGSPQGVIP 124
DB 64 QFDAYGDSATQREIYMGSVCHILPHLLIGONASVFAVGTGAGKTHMLGNPSQGVIP 123

QY 125 RALMDLLQLTREE--GAEGRPWALSVMYSLEYQEKVLDLDPASGDLVIREDCRGNIL 182
DB 124 RAVRDLLQMTTAAGPENENWTITMSYVEIYQEKVMDLLEPKNKDLPREDKDHNIL 183

QY 183 IGPLSQKPISSPADFERHFLPASNRVTGATRLNQRSSRSHAVLLVKVQDRELAPROR 242
DB 184 IPGVTQKTINSFGDFDEHFIPASQNTVASTKLNDRSSRSHAVLLIKVQSQVSPRQL 243

QY 243 EGKLYLIDLAGESENRRTGNKGLRKESGAINTSFLVLGKVVDALNQLPRVYRDSKLT 302
DB 244 TGKLYLIDLAGESENRRTGNKGLRKESGAINTSFLVLGKVVDALNQLPRVYRDSKLT 303

QY 303 RLLQDSLGSASHILIANIAPERFDYTSALNFAARKEVINRPTNSLQPHALGPV 362
DB 304 RLLQDSLGSASHILIANIAPERFDYTSALNFAARKEVINRPTNSLQPHALGPV 363

QY 363 KLSQKELLGPENKARPEEEIGSPPEMAAPASAKLSPLQK--LSSMDPAMLERLL 420
DB 364 MKRPRE-----EAETAAGSRQKKSKTSTESSNTMDAASKRLNLAALDPVAVRELL 418

QY 421 SLDRLLASQGSOCAPLLTPKRRMVLMTVEEKDLERLTKTKQLEAKQAKAE-E 479
DB 419 KLDKILTEGKKEAQLLSTPKRRMVLMTVEEKDLERLTKTKQLEAKAEAEARLE 478

QY 480 KENH 483
DB 479 KSTN 482

```

Db 304 RLQSLGSGTAHSVMIANIAPEQKYFDLTALNFAAKSKQIINKPFSQETTSQTAALPA 363
Qy 363 KLSQKELGPPPEAKRGPPEEEIGSPPEFMAAPASASQKLSPLQK--LSSMDPAMLERLL 420
Db 364 MKRPRE-----EAETAAGSRQKSKTSTESSPNTSDAASKRKLNALDPAVVERLL 418
Qy 421 SLDRLLASQSGQAPLLSTPKRVMVMTVEEKDLEIRLTKTKQKELEAKMLAQKAE-E 479
Db 419 KLDKILTEKGMKEAQLSTPKRERMALLKKWESQMEYERLKEKQKELEQKAEAEARLE 478
Qy 480 KENH 483
Db 479 KSTN 482

RESULT 6
Q91869 PRELIMINARY; PRT; 651 AA.
ID AC Q91869;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Kinesin (Hypothetical protein) (Chromokinesin Xkid).
GN Name=KID;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20419289; PubMed=10966105;
RA Antonio C., Ferby I., Wilhelm H., Jones M., Karsenti E., Nebreda A.R.,
RA Vernos I.;
RT "Xkid, a chromokinesin required for chromosome alignment on the
RT metaphase plate.";
RL Cell 102:425-435 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs S.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [4]
RP SEQUENCE FROM N.A.

```

```

RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Blastula;
RX MEDLINE=99340303; PubMed=10411507;
RA Zou H., McGarry T.J., Bernal T., Kirschner M.W.;
RT "Identification of a vertebrate sister-chromatid separation inhibitor
RT involved in transformation and tumorigenesis.";
RL Science 285:418-422 (1999).
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE=Blastula;
RX MEDLINE=20419288; PubMed=10966104;
RA Funabiki H., Murray A.W.;
RT "The Xenopus chromokinesin Xkid is essential for metaphase chromosome
RT alignment and must be degraded to allow anaphase chromosome
RT movement.";
RL Cell 102:411-424 (2000).
RN [7]
RP SEQUENCE FROM N.A.
RC TISSUE=Blastula;
RA McGarry T.J., Bernal T., Funabiki H., Kirschner M.W.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ249840; CAB71798.1; -
DR EMBL; BC070549; AAH70549.1; -
DR EMBL; AF267849; AAF26563.1; -
DR HSSP; P33173; 1155.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005875; C:microtubule associated complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR InterPro; IPR010996; DNaPol_B_N_like.
DR InterPro; IPR003583; HH1.1.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00278; HH1; 2.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
DR KW ATP-binding; Hypothetical protein; Microtubule; Motor protein.
SQ SEQUENCE 651 AA; 73065 MW; C2ABA561C4C53C13 CRC64;

Query Match 58.2%; Score 1438; DB 2; Length 651;
Best Local Similarity 60.1%; Pred. No. 1.1e-85;
Matches 292; Conservative 100; Mismatches 100; Indels 14; Gaps 7;

Qy 5 RLKIGATATPPPPARVAVRLRPFVDTAGA-SDPPCVRGMDSCSLEIANWRNHQETLK 63
Db 17 RVSMLDQHKSSCARVAVRLRPPYMDKDEAKATTCVVRGLDSQSLEIVNRRQLETMQ 76
Qy 64 YQFDAPFYGERSTQDDIYAGSVQPIRLHLLBQNASVLAYGPTGAGTKHTTLMGSPQPGVI 123
Db 77 YQFDAPFYGDSASQREIYMGSVCHLPHLLIGONASVAYGPTGAGTKHTTLMGNPQPGVI 136
Qy 124 PRALMDLLQUTREGA--EGRPWALSVMTSYLEYIQEKVLDLDPASGDLVIREDCGNI 181
Db 137 PRAVRDLLQMSRTAASAPENENWTYTNMSYVEIYQEKVMDLLEPKNKLPIREDKQHN 196
Qy 182 LIPGLSKPTSSPADPERHFLPASRNTVGTALNRSSRSASHVLLVKKVQDRELRAPRQ 241
Db 197 LIPGVTKMINSFADFEHFIPASQNTVASTKLNDRSSSHAVLLIKVQKSQVVPFRQ 256
Qy 242 REGKLYLDLAGSEDNRRTNGKLRKESGAINTSLFVLGKVVDALNQGLPRVPRYRDKL 301
Db 257 LTGKLYLDLAGSEDNERTNGQIRKESGAINTSLFVLGKVVDALNQGLPRIYRDKL 316
Qy 302 TRLLQSLGGSAAHSILIANIANTAPRRFYLDVTVAINFARKEVINRPTNESLQFHALGP 361
Db 317 TRLLQSLGGSAAHSVMTNTIAPEQTYFDLTALNFAAKSKQIINKPFSQETTSQTAALPA 376

```


DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.
SQ SEQUENCE 651 AA; 73091 MW; C4EF79801E603C13 CRC64;

Query Match 58.1%; Score 1437; DB 2; Length 651;
Best Local Similarity 60.1%; Pred. No. 1.3e-85;
Matches 292; Conservative 80; Mismatches 100; Indels 14; Gaps 7

Qy 5 RLSKIGATRRPPPARVRVAVRLRPFVDGTAGA--SDPPCVGRWMDSCSLSEIANRNHQETLK 63
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 17 RVSMULDQHKKSSCARVAVRLRPVMDEKDAKTTCVCRGLDSQSLEIVNWRQLTQM 76
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 64 YQDPAFGERSTQDDIYAGSVQPIRLHLLLEGONASVLAYGPTGAKTHTMLGSPEQPQVI 123
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 77 YQDPAFYGDSSAQREIYMGSVCYLPHLLIQNASVFAYGPTGAKTHTMLGNPNQPGVI 136
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 124 PRALMDDLQLTRREGA--EGRPWALSVTMSYLEIYOEKVLDLDPASGDLVIREDCRNI 181
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 137 PRAVRDLLQWSRTAAAPENWTYYTNMTSVEIYOEKVMDLLSPKNKDLPREDKHNI 196
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 182 LIPGLSKPISFFADFERHFPLPASNRITVGATRLNQSRSSRSASHVALLVKVDORERLAPP 241
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 197 LIPGVTKMINSFADFDEFIPAQNQRNVASTKLNDRSSRSASHVALLIKVKSQQVVPRQ 256
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 242 REGKLYIDLAGSEDNRRTNGKURLKESGAINTSLFVLGKVVDALNQGLRPVPYRDKL 301
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 257 LTGKLYIDLAGSDDNRRTNGQIRLKESGAINSLFTLSKVVDALNQGLPRIPYRDKL 316
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 302 TRLLQDSLGSGSAHSILNTANTAPERFYLDTVTSALNFARSKEVINRPTNESLQPHALGP 361
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 317 TRLLQDSLGSGSAHSVIMTNAPEQTYYFDTLTALNFAAKSKOIINKPSQETTQTVVQA 376
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 362 VKLSOKS---LLGPPEAKRGCPDEEEIGSGPEPMAAPASASOKLSPLQKLTSSMDPAMLER 418
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 377 MKRPRETGHIAAGSKRKSKNDSTE--SSPN-SMSTAGKQKLN----LATLPNAVVER 429
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 419 LLSDRLLASQSGOGAPLLSTPKRMVMLMHTVBCKOLEIERLKYTKQEELAKMLAQAE 478
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 430 LLKLDKITLKGKKKAQLLSTPKRMVALLKKWEESQMEIERLKEKQELFQKAMEAR 489
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 479 -KENH 483
|| || :
Db 490 LEKSN 495

RESULT 9
ID Q7ZYU5 PRELIMINARY; PRT; 663 AA.

AC Q7ZYU5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to kinesin family member 22 (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OX xenopodinae; Xenopus.
NCBI_Taxid=8355;
[1]
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
LA Klein S., Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC043733; AAA43733.1; -.
DR HSSP; P33173; IFS5.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005875; C:microtubule associated complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003774; P:motor activity; IEA.
DR InterPro; IPRO10996; DNaPol_B_like.
DR InterPro; IPRO03583; HH_1.
DR InterPro; IPRO01752; kinesin motor.

```

DR EMBL; AJ249841; CAB71799.1; -.
DR HSP; P31173; I155.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0008975; C:mitochondrion associated complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR InterPro; IPR010996; HNH1_B_N-like.
DR InterPro; IPR003583; HNH1.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR SMART; SM00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN2; 1.
DR PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.
SQ SEQUENCE 650 AA; 73027 MW; EA8415BC4B9B0F4A CRC64;

Query Match 57.6%; Score 1423; DB 2; Length 650;
Best Local Similarity 59.4%; Pred. No. 1e-84;
Matches 287; Conservative 81; Mismatches 107; Indels 8; Gaps 4;

QY 5 RLSKIGATRPAPRVAVRLPFDVGTAGASDP-PCVRGMDSCSLEIANRNHNETLQ 63
DB 17 RASMLDQHKSSCARVRVRLPYMEKEEDKVPACVRGLDSSHLEIVNRNQLTQ 76
QY 64 YQDFAFGERSTODIYAGSVQPIRLHLEGGQNASVLAYGPTGAGKTHMLGSPQPGVI 123
DB 77 YQDFAFGDSASQREIYWGVSCHILPHLLIGQNASVFAYGPTGAGKTHMLGPNQPGVI 136
QY 124 PRALMDLQLTR--EGAGRPWALSVTMSYLIYQEKVLDLDPASGLDVIREDCRGNI 181
DB 137 PRAVRELLQTRMAASAPENENWTHITTSYVEIYQEKVMDLLEPKNKDLPREDKDHNI 196
QY 182 LIPLGSKPISFADPEHFLPASRNTVGTATLNORSSSHAVLLVKVDORERLAPFRQ 241
DB 197 LIPGVTLTINSFGDFAHFIPASQNTVASTKLNDRSSSHAVLLIKVQSQVAPFRQ 256
QY 242 REGKYLIDLAGSDNRNRTGKGLRKESGAINSTLFLVKVVDALNQLGPRVPRYRDKL 301
DB 257 LIGKYLIDLAGSDNRNRTGKGLRKESGAINSTLFLVKVVDALNQLGPRVPRYRDKL 316
QY 302 TRLQDSLGSAHSILANTAPERFYLDTVSALNFAARKEVINPFTNESIQHALGP 361
DB 317 TRLQDSLGSAHSVMTNTIAPEQTYFDLTALNFAAKSKQILNKPFSEITQTVAQPA 376
QY 362 VKLSQKELLGPPEAKRGPEEIRIGSPPEMAAPASQKLSPLQKLSMDPAMLERLLS 421
DB 377 MKRPREARATTSRQRKSKTDSSTSPNSMESTGKRLN---LASLDSAVVERLLK 432
QY 422 LDRLLASQGSQCAPLLSTPKRVMVLMKTVEEKDLEIRLTKQKELEAKMLQAKAB-EK 480
DB 433 LDKILTEGKKEAQLLSTPKRVMVLMKTVEEKDLEIRLTKQKELEAKMLQAKAB-EK 492
QY 481 ENH 483
DB 493 SNN 495

RESULT 11
QSGPGO PRELIMINARY; PRT; 631 AA.
AC Q6GPGO;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;

```

```

RN RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haie H.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rask S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC073177; AAH73177.1; -.
DR InterPro; IPR001752; kinesin motor.
DR InterPro; IPR010994; Ruva_2_like.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Hypothetical protein; Microtubule; Motor protein.
SQ SEQUENCE 631 AA; 71084 MW; 1158C9AB86EBD3F3 CRC64;

Query Match 57.5%; Score 1422; DB 2; Length 631;
Best Local Similarity 60.6%; Pred. No. 1.2e-84;
Matches 285; Conservative 78; Mismatches 99; Indels 8; Gaps 4;

QY 18 ARVRVAVRLRPFDVGTAGASDP-PCVRGMDSCSLEIANRNHNETLKYQDFAFYGERSTQ 76
DB 11 ARVRVAVRLRPYMEKEEDKVPACVRGLDSSHLEIVNRNQLTQYQDFAFYGDASQ 70
QY 77 QDIYAGSVQPIRLHLEGGQNASVLAYGPTGAGKTHMLGSPQPGVIPALMDLQLTR- 135
DB 71 REIYWGVSCHILPHLLIGQNASVFAYGPTGAGKTHMLGNPDQPGVIPRAVRELLQTRM 130
QY 136 -EGAGRPWALSVTMSYLIYQEKVLDLDPASGLDVIREDCRGNILPGLSQKPISSF 194
DB 131 AASAPENENWTHITTSYVEIYQEKVMDLLEPKNKDLPREDKDHNLIPGVTLKTNF 190
QY 195 ADPERHFLPASRNTVGTATLNORSSSHAVLLVKVDORERLAPFRQREGKLYLIDLAGS 254
DB 191 GDFDEHFIPASQNTVASTKLNDRSSSHAVLLIKVQSQVAPFRQLGKLYLIDLAGS 250
QY 255 EDNRRTGNKGLRKESGAINSTLFLVKVVDALNQLGPRVPRYRDKSLTLLQDSLGSAH 314
DB 251 EDNRRTGNQIRLKESGAINSTLFLVKVVDALNQLGPRIPYRDKSLTLLQDSLGSAH 310

```



```

Qy 315 SILIANTIPERRFYLDVTSALNFAARKEVINRFTNESLQPHALGPVKLSQKELGPPE 374
Db 311 SVMITNIAPEQTYFDLTALNFAAKSKQIINKPFSRETTQVAQPAKMRPREAEATTS 370
Qy 375 AKRARGPEERIGSPPEWAPASAKLSPLQKLSMDPAMLERLLSLDRLLASQSGGA 434
Db 371 SRQKRSKTDSTESPNMSMESTGKRLN----LASLDSAVVERLLKLDKLTLEKGRKEA 426
Qy 435 PLLSTPKRERMLVMTVEEKLEIERLKTQKLEAKMLAQKAE-EXENH 483
Db 427 QLLSTPKRERMLVMTVEEKLEIERLKTQKLEAKMLAQKAE-EXENH 476

RESULT 12
O60845 PRELIMINARY; PRT; 198 AA.
AC O60845;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to kinesin-like DNA binding protein (KID).
GN Name=A-328A3.2;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=99425270; PubMed=10493829;
RA Loftus B.-J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,
RA Fuhrmann J., Mason T., Crosby M.B., Barnstead M., Cronin L.,
RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.L., Mitchell S.,
RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
RT "Genome duplications and other features in 12 Mb of DNA sequence from
RT human chromosome 16p and 16q.";
RL Genomics 60:295-308(1999).
DR EMBL; AC002301; AAC08709.1; -.
DR HSP; P33173; I161.
DR GO; GO:0005875; C:microtubule associated complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003774; P:motor activity; IEA.
DR InterPro; IPR001752; Kinesin_motor.
DR Pfam; PF00225; Kinesin; 1.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.
SQ SEQUENCE 198 AA; 21567 MW; 1A7FDA95B5A03B37 CRC64;

Query Match 33.8%; Score 836; DB 2; Length 198;
Best Local Similarity 100.0%; Pred. No. 5.7e-47;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRCLSKIGATRRPPPARVAVRLRPFVDTAGASDPPCVRGNDSCSLIANWRNHQET 61
Db 26 GRCLSKIGATRRPPPARVAVRLRPFVDTAGASDPPCVRGNDSCSLIANWRNHQET 85
Qy 62 LKQFDFAFYGERSTQDIYAGSVOPILRHLLGQNASVLYAGPTGACKTHMLGSPQPG 121
Db 86 LKQFDFAFYGERSTQDIYAGSVOPILRHLLGQNASVLYAGPTGACKTHMLGSPQPG 145

Qy 122 VIPRALMDLLQLTREGAAGRPWALSVTMSYLEIYQEKV 160
Db 146 VIPRALMDLLQLTREGAAGRPWALSVTMSYLEIYQEKV 184

RESULT 13
O35232 PRELIMINARY; PRT; 148 AA.
AC O35232;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Kinesin motor protein KIF22 (Fragment).

```

```

GN Name=Kif22;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=97480723; PubMed=9339368;
RA Yang Z., Hanlon D.W., Marszalek J.R., Goldstein L.S.;
RT "Identification, partial characterization, and genetic mapping of
RT kinesin-like protein genes in mouse.";
RL Genomics 45:123-131(1997).
DR EMBL; AF013119; AAC39968.1; -.
DR MGD; MGI:109233; Kif22.
DR GO; GO:0000785; C:chromatin; IEA.
DR GO; GO:0005819; C:spindle; IEA.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.
FT NON_TER 1
FT NON_TER 148
SQ SEQUENCE 148 AA; 16596 MW; 0950C265C3B63EB7 CRC64;

Query Match 28.5%; Score 704; DB 2; Length 148;
Best Local Similarity 92.6%; Pred. No. 1.7e-38;
Matches 137; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 108 GKTHMLGSPQCGVIPRALMDLLQLTREGAAGRPWALSVTMSYLEIYQEKVLDLDDPA 167
Db 1 GKTHMLGSPQCGVIPRALMDLLQLTREGAAGRPWALSVTMSYLEIYQEKVLDLDDPA 60
Qy 168 SGDLVIREDCRGNILIPGLSQPISSPADFERHPLPASRNTVGATRLNQRSSSHAVLL 227
Db 61 SGDLVIREDCRGNILIPGLTQKPIISFDEQHPFASRNRVAVGATRLNQRSSSHAVLL 120
Qy 228 VKVDQERLERLAPFRQREGKYLIDLAGSE 255
Db 121 VKVDQERLERLAPFRQREGKYLIDLAGSE 148

RESULT 14
O6L512 PRELIMINARY; PRT; 584 AA.
AC O6L512;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative kinesin-like DNA binding protein.
GN Name=OJ1281_H05.2;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
SEQUENCE FROM N.A.
RA Chow T.-Y., Hsing Y.-I.-C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC117265; AAT39162.1; -.
DR InterPro; IPR010996; DNAPol_B_N_like.
DR InterPro; IPR000445; HhH.
DR InterPro; IPR001752; Kinesin_motor.
DR Pfam; PF00633; HHH; 1.
DR Pfam; PF00225; Kinesin; 1.

```

```

Best Local Similarity 36.1%, Pred. No. 5.1e-33;      15;
Matches 166; Conservative 89; Mismatches 155; Indels 50; Gaps

QY      6 LSKIGATRRPPPARVAVVRLRPFVDGTAGASDPPPCVRGMDS-CSLEIANWRNHQETLK- 63
Db      1 MATAAATQSQP-----VRVTLVRPHLSEANSAPCVGLLGSHPGGEVTVQLKDQYTSRN 57

QY     64 --YQFDAGYERSTOODIYAGSVOPILRHLLLEGONASVLAYGPTGAGKTHTMLGSPQPG 121
Db      58 ECYKLDFAFGQESKVCIEIFDOEVSAVIPGIFEGTINATVFAYGATGSGKTYTMOQETEDLP 117

QY    122 VIPRALMDLLQLTREEGAGRPWALSVTMSVLEYIQSKVLDDLPASGDILVIREDCRGN 181
Db    118 LIPLAVSTVIALCT-----GTWC-SVEISYEVYEMERCVDLLEPKAREINVLDDKGNL 170

QY    182 LIPLSOKPTSSFADFERHELFPASRNTVGTATRLNQRSSRSHAVLLVKVDQERLAPFRQ 241
Db    171 QLKGLAWPVRSLEEFHEIYSIGVQRKVAHTGLNDVSSRSHAVLSIRITT-----DV 223

QY    242 REGKLYLIDLAGSEDNRRKTNGKGLRKESGAINTSLSFLVKGVDMALNQGLRPVRYRDKL 301
Db    224 VKGKLNLIIDLAGNEDNRRKTNEGIRLQESAKINOSLFALSNIIVSALNKKEPRIIYRESKL 283

QY    302 TRLLQDSIGSASHSILLTANTAPERFYLDTVSALNFAARSKEVINRPTNESLOPHALGP 361
Db    284 TRILQDSLGGNSHAWMLACLNPFVE--YOEAVHTVSLAARSRHVTNH-MSSASKQETPKDK 340

QY    362 VKLSOK-----ELLGPPPEA-KRARGPBEIEIGSPPEMAAP-----ASASQKSLPQKL 408
Db    341 VDMAEAKRLAWLESKGYKTSIQRMQD-----LLSPNAIKTPLSMGHKKQASGRVSGRGA 395

QY    409 SSMDPAMLERLL--SLDRLLAS-----QGSQGAPLLSTPKR 442
Db    396 MDQDGGIKKVLFFQSAARIIPAEINFHREGTQ--DIVNTTKK 433

```

Search completed: November 5, 2004, 18:38:43
Job time : 73.979 secs

```

PRINTS; PRO0380; KINESINHEAVY.
SMART; SMO0129; KISC; 1.
PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1.
ATP-binding; Microtubule; Motor protein.
SEQUENCE 584 AA; 64124 MW; EF829C3A15F9190F CRC64;

Query Match          25.6%; Score 633, DB 2; Length 584;
Best Local Similarity 36.1%; Pred. No. 5.1e-33;
Matches 166; Conservative 89; Mismatches 155; Indels 50; Gaps 15;

6  LSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPCVGRGMS- CSLEIANRNHNHOETLK- 63
1  MATAAATQSP---VRVLAVRPHLPSEANSAEAPCVGLLGHGPGGEVTVQLKDQYTSRN 57
64 --YQDFAFYGERSTQDDIYAGSVQPIRLHLLSQNASVLAYGPTGAGKTHTMLGSPQPG 121
58 ECYKLDIAFFGQBSRVCSIFQGEVASVPGIFEGTNAVTFAYGATSGGKTYTNQGTEDLPG 117
122 VIPRALMDLLQLTREEAGCRPWALSVTMSYLEIYQEKVLDLIDPASGLDVIREDRCNGI 181
118 LIPLAVSTVLAICT-----GTWC-SVEISVYEYVMERCYDLELPEKAREIMVLDDKGNL 170
182 LIPGLSQKPISSFADFERHFLPASRNTVTGATLNRSSRSRAVLVKKVDQERLAPRQ 241
171 QLKGLAVPVRSLBEFHEIYSIGVQRKVAHTGLNDVSSRSRAVLSIRITT-----DV 223
242 REGKLYLDLAGSDNDRNTGNKGLRKESAINTSFLVLGKVVADALNOGLPRVPYRDSKL 301
224 VQKLNLLIDLAGNEDNRTCNEGIRLQESAKINQSLFALSNVISALNKKEPAIPYRESKL 283
302 TRLLODSLGGSAHSILIANIAPERFRFYDVTVSALNFAARSKVINPPTNESLOPHALGP 361
284 TRILQDSLGGNSHAWIACLNPVE--YQEAHTVTSLAARSRHYTNH-MSASQKETPKDK 340
362 VKLSQK-----ELLGPEEA-KRARGPEEEIGSPPEMAAP-----ASASQKLSPLQKL 408
341 VDMERAKLRAWLESKTKYSKIORMDG-----LLSPNAIKTPLSMHKKQASGRVSGRGA 395
409 SSMDDPAMLERLL--SLDRLLAS----QGSQAPLLSTPKR 442
396 MPODGGKTKVLFDSAAIRIPANTFHREGTO--DIVNTTKK 433

```

Query Match

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 ; Search time 20.0023 Seconds
(without alignments)
1697.543 Million cell updates/sec

Title: US-10-797-893-4

Perfect score: 2589

Sequence: 1 MPAAGGSGTQRRREMAAAS.....LEAKMLAQKAEKENHCPM 512

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/6C_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2589	100.0	512	3	US-09-724-224-4
2	2589	100.0	512	4	US-10-093-317-4
3	2577	99.5	665	4	US-09-595-684B-35
4	2467	95.3	487	3	US-09-724-224-8
5	2467	95.3	487	4	US-10-093-317-8
6	1873	72.3	370	3	US-09-724-224-2
7	1873	72.3	370	4	US-10-093-317-2
8	1763	68.1	346	3	US-09-724-224-6
9	1763	68.1	346	4	US-10-093-317-6
10	573.5	22.2	1234	4	US-09-592-054-8
11	568.5	22.0	473	4	US-09-592-054-6
12	568.5	22.0	522	4	US-09-592-054-4
13	561.5	21.7	1232	4	US-09-592-054-2
14	561	21.7	1388	3	US-09-572-191-2
15	561	21.7	1388	3	US-09-723-262-2
16	561	21.7	1388	3	US-09-723-219-2
17	556.5	21.5	1066	3	US-09-541-782-8
18	556.5	21.5	1066	4	US-09-723-820-8
19	556.5	21.5	1066	4	US-10-270-085-8
20	554	21.4	1279	3	US-09-724-517-2
21	554	21.4	1279	4	US-09-641-807A-2
22	554	21.4	1279	4	US-09-723-096-2
23	553.5	21.4	1231	4	US-09-595-684B-23
24	551	21.3	2954	4	US-09-150-867-1
25	550	21.2	864	4	US-09-883-096-2
26	549	21.2	355	3	US-09-724-511-4
27	549	21.2	355	4	US-09-723-097-4

28	549	21.2	355	4	US-09-632-344-4	Sequence 4, Appli
29	549	21.2	367	3	US-09-724-511-2	Sequence 2, Appli
30	549	21.2	367	4	US-09-723-097-2	Sequence 2, Appli
31	549	21.2	367	4	US-09-632-344-2	Sequence 2, Appli
32	546.5	21.1	1103	3	US-09-162-373-1	Sequence 1, Appli
33	546.5	21.1	1103	3	US-09-467-946-1	Sequence 25, Appli
34	544.5	21.0	935	4	US-09-914-259-25	Sequence 6, Appli
35	543	21.0	409	3	US-09-572-191-6	Sequence 6, Appli
36	543	21.0	409	3	US-09-723-262-6	Sequence 6, Appli
37	543	21.0	409	3	US-09-723-219-6	Sequence 23, Appli
38	543	21.0	928	4	US-09-914-259-23	Sequence 39, Appli
39	541	20.9	1690	4	US-09-595-684B-39	Sequence 6, Appli
40	535	20.7	513	4	US-09-724-519-6	Sequence 6, Appli
41	535	20.7	513	4	US-09-592-037-6	Sequence 6, Appli
42	535	20.7	513	4	US-09-428-156B-6	Sequence 2, Appli
43	533.5	20.6	1057	4	US-09-428-156B-2	Sequence 29, Appli
44	532	20.5	1056	4	US-09-595-684B-29	Sequence 10, Appli
45	532	20.5	1057	3	US-09-541-782-10	

ALIGNMENTS

RESULT 1
US-09-724-224-4
; Sequence 4, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09/724, 224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597, 292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Human
US-09-724-224-4

Query Match		100.0%;	Score 2589;	DB 3;	Length 512;
Best Local Similarity		100.0%;	Pred. No. 1.8e-233;		
Matches 512;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MPAAGGSGTQRRREMAAASAAAI	SGAGRCRLSKIGATRRPPPARVAVRLRPPFVDGTAG	60	
Db	1	MPAAGGSGTQRRREMAAASAAAI	SGAGRCRLSKIGATRRPPPARVAVRLRPPFVDGTAG	60	
Qy	61	ASDPCCVRGMDSCSLEIANWRNHQETLKYQF	DAFYGERSTQDDIYAGSVQPIRLHLLGGQ	120	
Db	61	ASDPCCVRGMDSCSLEIANWRNHQETLKYQF	DAFYGERSTQDDIYAGSVQPIRLHLLGGQ	120	
Qy	121	NASVLAYGPTGAGKTHMLGSPQPGV	IPRALMDLLQITREEGAGRPWALSVTMSYLEI	180	
Db	121	NASVLAYGPTGAGKTHMLGSPQPGV	IPRALMDLLQITREEGAGRPWALSVTMSYLEI	180	
Qy	181	YQEKVLDLLDPASGDLVIREDCRGNILIP	GLSQPISSFADFERRHFLPASNNRTVGATRL	240	
Db	181	YQEKVLDLLDPASGDLVIREDCRGNILIP	GLSQPISSFADFERRHFLPASNNRTVGATRL	240	
Qy	241	NQRRSRSHAVLLVKVDQERLAPFRQREGK	LYLIDLAGSEDNRRRTGNKGLKESGAINT	300	
Db	241	NQRRSRSHAVLLVKVDQERLAPFRQREGK	LYLIDLAGSEDNRRRTGNKGLKESGAINT	300	
Qy	301	SLFVLGKVDALNOGLPRVYRDSKLT	RLLODSLGGSAAHSILIANIAPERRFYLDTSAL	360	
Db	301	SLFVLGKVDALNOGLPRVYRDSKLT	RLLODSLGGSAAHSILIANIAPERRFYLDTSAL	360	
Qy	361	NFAARSEVINRPTNESLQPHALGPVKLSQ	KELLGPPEAKRGPEEEIGSPPEPMAAP	420	

Db 361 NFAARKEVINRPTNESLOPHALGPVKLSQKELLGPPEAKRGPEEEIIGSPPEWAP 420
Qy 421 ASASQKLSPLQKLSMDPAMLERLLSLDRLLASQSGAPLLSTPKRERVMVLMKTVEEKD 480
Db 421 ASASQKLSPLQKLSMDPAMLERLLSLDRLLASQSGAPLLSTPKRERVMVLMKTVEEKD 480
Qy 481 LEIERLTKQKELEAKMLAQKAEKENHCPTM 512
Db 481 LEIERLTKQKELEAKMLAQKAEKENHCPTM 512

RESULT 2
US-10-093-317-4
; Sequence 4, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093,317
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Human
US-10-093-317-4

Query Match 100.0%; Score 2589; DB 4; Length 512;
Best Local Similarity 100.0%; Pred. No. 1.8e-233;
Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPAGGSTQORREMAAASAAISGAGRCRLSKIGATRRPPARVAVLRPFVDTAG 60
Db 1 MPAGGSTQORREMAAASAAISGAGRCRLSKIGATRRPPARVAVLRPFVDTAG 60
Qy 61 ASDPPCVRGMDSCSLEIANWNHQETLKYQDFAFYGERSTQODIYAGSVQPIRLHLLGQ 120
Db 61 ASDPPCVRGMDSCSLEIANWNHQETLKYQDFAFYGERSTQODIYAGSVQPIRLHLLGQ 120
Qy 121 NASVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTREBAGRPPWALSVTMSYLEI 180
Db 121 NASVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTREBAGRPPWALSVTMSYLEI 180
Qy 181 YQSKVLDLLDPASGDLVIREDCRGNILIPGLSOKPISSFADPERHFLPASRNTVGTATL 240
Db 181 YQSKVLDLLDPASGDLVIREDCRGNILIPGLSOKPISSFADPERHFLPASRNTVGTATL 240
Qy 241 NQSSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGSEDNRRRTGNKGLRKESGAIN 300
Db 241 NQSSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGSEDNRRRTGNKGLRKESGAIN 300
Qy 301 SLFVLGVVDALNOGLPRVPYRDSKLTRELQDSLGSSAHSILIANIAPERFYLDTVSAL 360
Db 301 SLFVLGVVDALNOGLPRVPYRDSKLTRELQDSLGSSAHSILIANIAPERFYLDTVSAL 360
Qy 361 NFAARKEVINRPTNESLOPHALGPVKLSQKELLGPPEAKRGPEEEIIGSPPEWAP 420
Db 361 NFAARKEVINRPTNESLOPHALGPVKLSQKELLGPPEAKRGPEEEIIGSPPEWAP 420
Qy 421 ASASQKLSPLQKLSMDPAMLERLLSLDRLLASQSGAPLLSTPKRERVMVLMKTVEEKD 480
Db 421 ASASQKLSPLQKLSMDPAMLERLLSLDRLLASQSGAPLLSTPKRERVMVLMKTVEEKD 480
Qy 481 LEIERLTKQKELEAKMLAQKAEKENHCPTM 512
Db 481 LEIERLTKQKELEAKMLAQKAEKENHCPTM 512

RESULT 3
US-09-595-684B-35
; Sequence 35, Application US/09595684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Ohashi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; TITLE OF INVENTION: and purifying human kinesins
; FILE REFERENCE: cytop036
; CURRENT APPLICATION NUMBER: US/09/595,684B
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Human
US-09-595-684B-35

Query Match 99.5%; Score 2577; DB 4; Length 665;
Best Local Similarity 100.0%; Pred. No. 3.7e-232;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AAGGSTQORREMAAASAAISGAGRCRLSKIGATRRPPARVAVLRPFVDTAGS 62
Db 2 AAGGSTQORREMAAASAAISGAGRCRLSKIGATRRPPARVAVLRPFVDTAGS 61
Qy 63 DPPCVRGMDSCSLEIANWNHQETLKYQDFAFYGERSTQODIYAGSVQPIRLHLLGQNA 122
Db 62 DPPCVRGMDSCSLEIANWNHQETLKYQDFAFYGERSTQODIYAGSVQPIRLHLLGQNA 121
Qy 123 SVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTREBAGRPPWALSVTMSYLEIYQ 182
Db 122 SVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTREBAGRPPWALSVTMSYLEIYQ 181
Qy 183 EKVLDDLLDPASGDLVIREDCRGNILIPGLSOKPISSFADPERHFLPASRNTVGTATLQ 242
Db 182 EKVLDDLLDPASGDLVIREDCRGNILIPGLSOKPISSFADPERHFLPASRNTVGTATLQ 241
Qy 243 RSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGSEDNRRRTGNKGLRKESGAIN 302
Db 242 RSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGSEDNRRRTGNKGLRKESGAIN 301
Qy 303 FVLGVVDALNOGLPRVPYRDSKLTRELQDSLGSSAHSILIANIAPERFYLDTVSALNF 362
Db 302 FVLGVVDALNOGLPRVPYRDSKLTRELQDSLGSSAHSILIANIAPERFYLDTVSALNF 361
Qy 363 AARSKEVINRPTNESLOPHALGPVKLSQKELLGPPEAKRGPEEEIIGSPPEMAAPAS 422
Db 362 AARSKEVINRPTNESLOPHALGPVKLSQKELLGPPEAKRGPEEEIIGSPPEMAAPAS 421
Qy 423 ASQKLSPLQKLSMDPAMLERLLSLDRLLASQSGAPLLSTPKRERVMVLMKTVEEKDLE 482
Db 422 ASQKLSPLQKLSMDPAMLERLLSLDRLLASQSGAPLLSTPKRERVMVLMKTVEEKDLE 481
Qy 483 IERLTKQKELEAKMLAQKAEKENHCPTM 512
Db 482 IERLTKQKELEAKMLAQKAEKENHCPTM 511

RESULT 4
US-09-724-224-8
; Sequence 8, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:

; APPLICANT: Beraud, Christophe
 ; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
 ; FILE REFERENCE: their use
 ; CURRENT APPLICATION NUMBER: US/09/724,224
 ; CURRENT FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: 09/597,292
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 487
 ; TYPE: PRT
 ; ORGANISM: Human
 ; US-09-724-224-8

Query Match 95.3%; Score 2467; DB 3; Length 487;
 Best Local Similarity 100.0%; Pred. No. 4.4e-222;
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 27 GRCRLSKIGATRRPPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANRNHOET 86
 Db 2 GRCRLSKIGATRRPPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANRNHOET 61
 Qy 87 LKQFDAPYGRSTQDDIYAGSVQPIRLHLLLEGQNASVLAIGPTGAGKTHMLGSPQPG 146
 Db 62 LKQFDAPYGRSTQDDIYAGSVQPIRLHLLLEGQNASVLAIGPTGAGKTHMLGSPQPG 121
 Qy 147 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 206
 Db 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 181
 Qy 207 LIPGLSKPISSPADFERHFLPASNRVTGATRLNQRSSSRSHAVLLVKVDQERLAPPRQ 266
 Db 182 LIPGLSKPISSPADFERHFLPASNRVTGATRLNQRSSSRSHAVLLVKVDQERLAPPRQ 241
 Qy 267 REGKLYLIDLAGSDNRRTGNKGLRKESGAINSTSLFVLGKVDALNQGLPRVPYRDSKL 326
 Db 242 REGKLYLIDLAGSDNRRTGNKGLRKESGAINSTSLFVLGKVDALNQGLPRVPYRDSKL 301
 Qy 327 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPFPTNESLOPHALGP 386
 Db 302 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPFPTNESLOPHALGP 361
 Qy 387 VKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKSSMDPAMLERLLS 446
 Db 362 VKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKSSMDPAMLERLLS 421
 Qy 447 LDRLLASQGSQCAPLLSTPKRERVMKTVEEKDLEIRLTKKQLEAKMLAKAEKE 506
 Db 422 LDRLLASQGSQCAPLLSTPKRERVMKTVEEKDLEIRLTKKQLEAKMLAKAEKE 481
 Qy 507 NHCPMT 512
 Db 482 NHCPMT 487

RESULT 5
 US-10-093-317-8
 ; Sequence 8, Application US/10093317
 ; Patent No. 6762043
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; TITLE OF INVENTION: their use
 ; FILE REFERENCE: 1044
 ; CURRENT APPLICATION NUMBER: US/10/093,317
 ; CURRENT FILING DATE: 2002-03-06
 ; PRIOR APPLICATION NUMBER: 09/724,224
 ; PRIOR FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8

; LENGTH: 487
 ; TYPE: PRT
 ; ORGANISM: Human
 ; US-10-093-317-8
 Query Match 95.3%; Score 2467; DB 4; Length 487;
 Best Local Similarity 100.0%; Pred. No. 4.4e-222;
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 27 GRCRLSKIGATRRPPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANRNHOET 86
 Db 2 GRCRLSKIGATRRPPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANRNHOET 61
 Qy 87 LKQFDAPYGRSTQDDIYAGSVQPIRLHLLLEGQNASVLAIGPTGAGKTHMLGSPQPG 146
 Db 62 LKQFDAPYGRSTQDDIYAGSVQPIRLHLLLEGQNASVLAIGPTGAGKTHMLGSPQPG 121
 Qy 147 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 206
 Db 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 181
 Qy 207 LIPGLSKPISSPADFERHFLPASNRVTGATRLNQRSSSRSHAVLLVKVDQERLAPPRQ 266
 Db 182 LIPGLSKPISSPADFERHFLPASNRVTGATRLNQRSSSRSHAVLLVKVDQERLAPPRQ 241
 Qy 267 REGKLYLIDLAGSDNRRTGNKGLRKESGAINSTSLFVLGKVDALNQGLPRVPYRDSKL 326
 Db 242 REGKLYLIDLAGSDNRRTGNKGLRKESGAINSTSLFVLGKVDALNQGLPRVPYRDSKL 301
 Qy 327 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPFPTNESLOPHALGP 386
 Db 302 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPFPTNESLOPHALGP 361
 Qy 387 VKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKSSMDPAMLERLLS 446
 Db 362 VKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKSSMDPAMLERLLS 421
 Qy 447 LDRLLASQGSQCAPLLSTPKRERVMKTVEEKDLEIRLTKKQLEAKMLAKAEKE 506
 Db 422 LDRLLASQGSQCAPLLSTPKRERVMKTVEEKDLEIRLTKKQLEAKMLAKAEKE 481
 Qy 507 NHCPMT 512
 Db 482 NHCPMT 487

RESULT 6
 US-09-724-224-2
 ; Sequence 2, Application US/09724224
 ; Patent No. 6387644
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
 ; FILE REFERENCE: their use
 ; CURRENT APPLICATION NUMBER: US/09/724,224
 ; CURRENT FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: 09/597,292
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 370
 ; TYPE: PRT
 ; ORGANISM: Human
 ; US-09-724-224-2

Query Match 72.3%; Score 1873; DB 3; Length 370;
 Best Local Similarity 100.0%; Pred. No. 1.1e-166;
 Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 3 AAGGSGTQORRREMAAASAAISGAGRCRLSKIGATRRPPPPARVRVAVRLRPFVDGTAGAS 62

Db 2 AAGGTOORREMAAASAAISGAGCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGAS 61
QY 63 DPPCVRGMDSCSLEIANWRNHQETLKYQDFAGYGERSTQDDIYAGSVQPIRLHLLGGQNA 122
Db 62 DPPCVRGMDSCSLEIANWRNHQETLKYQDFAGYGERSTQDDIYAGSVQPIRLHLLGGQNA 121
QY 123 SVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREBAGRPPWALSVTMSYLEIYQ 182
Db 122 SVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREBAGRPPWALSVTMSYLEIYQ 181
QY 183 EKVLDDLPASGDLVIREDCRGNILPGLSQKPISSFADFPERHFLPASRNRVTGATRLNQ 242
Db 182 EKVLDDLPASGDLVIREDCRGNILPGLSQKPISSFADFPERHFLPASRNRVTGATRLNQ 241
QY 243 RSSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGESEDNRRTGNKGLRKSAGINTSL 302
Db 242 RSSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGESEDNRRTGNKGLRKSAGINTSL 301
QY 303 FVLGKVDALNOGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERRFYLDTVSALNF 362
Db 302 FVLGKVDALNOGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERRFYLDTVSALNF 361
QY 363 AARSKEVIN 371
Db 362 AARSKEVIN 370
RESULT 7
US-10-093-317-2
; Sequence 2, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093,317
; PRIOR FILING DATE: 2002-03-06
; PRIOR FILING DATE: 09/724,224
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Human
US-10-093-317-2
Query Match 72.3%; Score 1873; DB 4; Length 370;
Best Local Similarity 100.0%; Pred. No. 1.1e-166; Indels 0; Gaps 0;
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 AAGGTOORREMAAASAAISGAGCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGAS 62
Db 2 AAGGTOORREMAAASAAISGAGCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGAS 61
QY 63 DPPCVRGMDSCSLEIANWRNHQETLKYQDFAGYGERSTQDDIYAGSVQPIRLHLLGGQNA 122
Db 62 DPPCVRGMDSCSLEIANWRNHQETLKYQDFAGYGERSTQDDIYAGSVQPIRLHLLGGQNA 121
QY 123 SVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREBAGRPPWALSVTMSYLEIYQ 182
Db 122 SVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREBAGRPPWALSVTMSYLEIYQ 181
QY 183 EKVLDDLPASGDLVIREDCRGNILPGLSQKPISSFADFPERHFLPASRNRVTGATRLNQ 242
Db 182 EKVLDDLPASGDLVIREDCRGNILPGLSQKPISSFADFPERHFLPASRNRVTGATRLNQ 241
QY 243 RSSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGESEDNRRTGNKGLRKSAGINTSL 302
Db 242 RSSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGESEDNRRTGNKGLRKSAGINTSL 301
QY 303 FVLGKVDALNOGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERRFYLDTVSALNF 362

Db 302 FVLGKVDALNOGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERRFYLDTVSALNF 361
QY 363 AARSKEVIN 371
Db 362 AARSKEVIN 370
RESULT 8
US-09-724-224-6
; Sequence 6, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR FILING DATE: 09/597,292
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Human
US-09-724-224-6
Query Match 68.1%; Score 1763; DB 3; Length 346;
Best Local Similarity 100.0%; Pred. No. 2e-156; Indels 0; Gaps 0;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 27 GRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 86
Db 2 GRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61
QY 87 LKYPDAFYGERSTQDDIYAGSVQPIRLHLLGGQNASVLAYGPTGAGKTHMLGSPQPG 146
Db 62 LKYPDAFYGERSTQDDIYAGSVQPIRLHLLGGQNASVLAYGPTGAGKTHMLGSPQPG 121
QY 147 VIPRALMDLLQLTREBAGRPPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRNI 206
Db 122 VIPRALMDLLQLTREBAGRPPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRNI 181
QY 207 LIPGLSQKPISSFADFPERHFLPASRNRVTGATRLNQRSSRSHAVLLVKVDQERLAPFRQ 266
Db 182 LIPGLSQKPISSFADFPERHFLPASRNRVTGATRLNQRSSRSHAVLLVKVDQERLAPFRQ 241
QY 267 REGKLYLIDLAGESEDNRRTGNKGLRKSAGINTSLFVLGKVDALNOGLPRVPYRDSKL 326
Db 242 REGKLYLIDLAGESEDNRRTGNKGLRKSAGINTSLFVLGKVDALNOGLPRVPYRDSKL 301
QY 327 TRLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 371
Db 302 TRLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346
RESULT 9
US-10-093-317-6
; Sequence 6, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093,317
; PRIOR FILING DATE: 2002-03-06
; PRIOR FILING DATE: 09/724,224
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Human
US-10-093-317-6

Query Match      68.1%; Score 1763; DB 4; Length 346;
Best Local Similarity 100.0%; Pred. No. 2e-156;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 GRCLSKIGATRRPPPARVRVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHOET 86
Db 2 GRCLSKIGATRRPPPARVRVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHOET 61

QY 87 LKYQFDAPYGRSTQODIYAGSVOPILRHLEGGONASVLAYGPTGAGKTHMLGSPGPG 146
Db 62 LKYQFDAPYGRSTQODIYAGSVOPILRHLEGGONASVLAYGPTGAGKTHMLGSPGPG 121

QY 147 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 206
Db 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 181

QY 207 LIPGLSQPISSPADFERHFLPASNRRTVGATRLNORSSRSHAVLLVKVDORERLAPRQ 266
Db 182 LIPGLSQPISSPADFERHFLPASNRRTVGATRLNORSSRSHAVLLVKVDORERLAPRQ 241

QY 267 REGKYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGVKVDALNQGRLPRVPYRDSKL 326
Db 242 REGKYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGVKVDALNQGRLPRVPYRDSKL 301

QY 327 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 371
Db 302 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346

RESULT 10
US-09-592-054-8
; Sequence 8, Application US/09592054
; Patent No. 6440684
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Finer, Jeffrey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1016
; CURRENT APPLICATION NUMBER: US/09/592.054
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1234
; TYPE: PRT
; ORGANISM: Human
US-09-592-054-8

Query Match      22.2%; Score 573.5; DB 4; Length 1234;
Best Local Similarity 32.2%; Pred. No. 3e-44;
Matches 156; Conservative 97; Mismatches 173; Indels 59; Gaps 13;

QY 45 VRVAVRLRPFDGTAGASDPPC-----VRGMDSCSLEIANWRNHOETLKYOPDAP 94
Db 10 VRVALRCPLVPKEISGQCMCLSFVPGTQVVVGTDK-----SFTYDFV 54

QY 95 YGERSTQODIYAGSVQPIRLHLEGGONASVLAYGPTGAGKTHMLGSPGPG--GV 147
Db 55 FDPCTEQUEEVFNKAVAPLIKGIFGYNATVLAYGQTGSGKTYSGGAYTAQEINEPTVI 114

QY 148 IPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 205
Db 115 IPR-----VIQLFKEDIKSDPFTLKVSYLYNEEILDLCPSREKAQINIREDPKEG 170

; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Human
US-10-093-317-6

Query Match      68.1%; Score 1763; DB 4; Length 346;
Best Local Similarity 100.0%; Pred. No. 2e-156;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 GRCLSKIGATRRPPPARVRVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHOET 86
Db 2 GRCLSKIGATRRPPPARVRVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHOET 61

QY 87 LKYQFDAPYGRSTQODIYAGSVOPILRHLEGGONASVLAYGPTGAGKTHMLGSPGPG 146
Db 62 LKYQFDAPYGRSTQODIYAGSVOPILRHLEGGONASVLAYGPTGAGKTHMLGSPGPG 121

QY 147 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 206
Db 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 181

QY 207 LIPGLSQPISSPADFERHFLPASNRRTVGATRLNORSSRSHAVLLVKVDORERLAPRQ 266
Db 182 LIPGLSQPISSPADFERHFLPASNRRTVGATRLNORSSRSHAVLLVKVDORERLAPRQ 241

QY 267 REGKYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGVKVDALNQGRLPRVPYRDSKL 326
Db 242 REGKYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGVKVDALNQGRLPRVPYRDSKL 301

QY 327 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 371
Db 302 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346

RESULT 11
US-09-592-054-6
; Sequence 6, Application US/09592054
; Patent No. 6440684
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Finer, Jeffrey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6440684el motor proteins and methods for
; FILE REFERENCE: 1016
; CURRENT APPLICATION NUMBER: US/09/592.054
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Human
US-09-592-054-6

Query Match      22.0%; Score 568.5; DB 4; Length 473;
Best Local Similarity 32.3%; Pred. No. 1.9e-44;
Matches 159; Conservative 97; Mismatches 162; Indels 75; Gaps 15;

QY 45 VRVAVRLRP-----FVDTGATAGSDPPCVRGMDSCSLEIANWRNHOETLKY 89
Db 8 VRVALRCPLVPKEISGQCMCLSFVPG-----EPQVVVGTDK-----SF 47

QY 90 QFDAFYGERSTQODIYAGSVQPIRLHLEGGONASVLAYGPTGAGKTHMLGSPGPG 144
Db 48 TYDFVDFDPTSEQBEVENTAVAPLIKGIFGYNATVLAYGQTGSGKTYSGGAYTAQEINE 107

QY 145 P--GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIRE 200
Db 108 PTGVGVIPIR-----VIQLFKEDIKSDPFTLKVSYLYNEEILDLCPSREKAQINIRE 163

QY 201 DCGNLIIPGLSQPISSPADFERHFLPASNRRTVGATRLNORSSRSHAVLLVKVDORER 260
Db 164 DPKEGKIVGLTEKTVLVALDVTSCLEQGNNSRTVASTAMNSOSSSRSHAFITISLEQKK 223

QY 261 LAFPRQREGKYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGVKVDALNQGRLPR 318
Db 224 SDKNSFRSKLHLVDLAGSERQKTKAEGRLKEGININRGLCLGNVISALGDDKKGKF 283

QY 319 VPYRDSKLRLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPFPTNES 378
```



```

; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6355466e1 motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1017
; CURRENT APPLICATION NUMBER: US/09/572,191
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Human
US-09-572-191-2

```

[illegible]

RESULT 15
US-09-723-262-2
; Sequence 2, Application US/09723262
; Patent No. 6379912
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6379912el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1017
; CURRENT APPLICATION NUMBER: US/09/723.262
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US 09/572,191
; PRIOR FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 6

[illegible]

Search completed: November 5, 2004, 18:46:45
Job time : 22.0023 secs

This page blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 ; Search time 105.087 Seconds
(without alignments)
1747.786 Million cell updates/sec

Title: US-10-797-893-4

Perfect score: 2589

Sequence: 1 MPAGSGTQORREMAAAS.....LEAKWLAQKAEEKNHCPPTM 512

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing, first 45 summaries

Database : A Geneseq_23Sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2589	100.0	512	7	ADC23340 Human kin
2	2589	100.0	512	8	ADQ60230 Human mic
3	2577	99.5	665	8	ADQ09240 Human KNS
4	2467	95.3	487	7	ADC23344 Human kin
5	2467	95.3	487	8	ADQ60234 Human mic
6	2419	93.4	490	7	ADK40973 Novel hum
7	2151	83.1	460	3	RAB56650 Human pro
8	1873	72.3	370	7	ADC23338 Human kin
9	1873	72.3	370	8	ADQ60228 Human mic
10	1763	68.1	346	7	ADC23342 Human kin
11	1763	68.1	346	8	ADQ60232 Human mic
12	616	23.8	784	4	ABB71112 Drosophil
13	593	22.9	677	4	ABB65183 Drosophil
14	578.5	22.3	1237	7	ADK40926 Novel hum
15	574.5	22.2	1232	7	ADK48938 Human lun
16	574.5	22.2	1232	7	ADD18924 Human dis
17	574.5	22.2	1232	7	ADJ70347 Human hea
18	574.5	22.2	1235	7	ADK40901 Novel hum
19	573.5	22.2	1234	5	ABG70993 Human HsK
20	568.5	22.0	473	5	ABG70992 Human tar
21	568.5	22.0	522	5	ABG70991 Human HsK
22	568	21.9	898	4	ABU53123 Intracell
23	568	21.9	898	5	ABU76957 Novel hum
24	568	21.9	898	7	ADC31082 Human nov
25	568	21.9	898	7	ADI15915 Human PP

26	567.5	21.9	998	8	ADO44167 Structura
27	564.5	21.8	548	7	ADM04007 Human pro
28	564.5	21.8	905	5	Aau76967 Novel hum
29	561.5	21.7	1232	5	ABg70990 Human HsK
30	561	21.7	1388	5	AAE14400 Human kin
31	561	21.7	1388	5	AAU79590 Human kin
32	561	21.7	1388	6	ABR48222 Human bla
33	561	21.7	1388	7	ADB80468 Ovarian c
34	561	21.7	1388	7	ADC35116 Human bre
35	561	21.7	1388	8	ADL83290 Human PRO
36	561	21.7	1388	8	ADQ20128 Human sof
37	561	21.7	1388	8	ADQ09226 Human KNS
38	560	21.6	972	8	ADN40548 Candida a
39	560	21.6	974	8	ADN40552 Candida a
40	559.5	21.6	1401	7	ADJ94914 Novel NOV
41	557.5	21.5	672	7	ADJ69858 Human hea
42	557.5	21.5	1232	7	ADD49937 Human kin
43	557.5	21.5	1232	7	ADD49932 Human lun
44	557	21.5	1324	7	ADJ95086 Novel NOV
45	556.5	21.5	1066	4	AAG67418 Amino aci

ALIGNMENTS

RESULT 1

ADC23340
ID ADC23340 standard; protein; 512 AA.

XX AC ADC23340;

DT 18-DEC-2003 (first entry)

DE Human kinesin-like DNA binding protein (KID) (SeqID 4).

XX human; enzyme; motor domain; kinesin-like DNA binding protein; KID;
KW cytosolic; cardiac; immunomodulator; antiinflammatory; gene therapy;
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;
KW cardiac hypertrophy; immune disorder; inflammation.

XX Homo sapiens.

FH Key Location/Qualifiers

FT Misc-difference 2 /note= "Encoded by CA"

XX US6387644-B1.

PD 14-MAY-2002.

PF 28-NOV-2000; 2000US-00724224.

PR 20-APR-1999; 99US-00295612.

PR 20-JUN-2000; 2000US-00597292.

XX (CYTO-) CYTOKINETICS INC.

XX Beraud C;

XX WPI; 2003-706919/67.

DR N-PSDB; ADC23339.

XX Identifying a candidate agent as modulator of function of a target
PT protein for treating cellular proliferation disorders by adding a
PT candidate agent to a mixture of the target protein that
PT directly/indirectly produces ADP or phosphate.

PS Claim 1; SEQ ID NO 4; 26pp; English.

XX This invention relates to a novel method for high throughput screening
CC systems used to identify compounds for the treatment of cellular
CC proliferation disorders. Specifically, it refers to candidate agents that
CC are capable of modulating the activity of target proteins having motor

CC domains, such that the target protein directly or indirectly produces ADP
CC or phosphate. Furthermore, this activity can be determined using
CC fluorescence or absorbance readouts. The present invention describes a
CC method that identifies modulators of the target protein, which is a
CC kinsin-like DNA binding protein (known as KID) as cytostatic, cardants,
CC immunomodulators and antiinflammatory. Accordingly, through gene
CC therapy, they can be used for the treatment of cancer, hyperplasias,
CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This
CC polypeptide sequence is human KID protein (SeqID 4) of the invention.
XX
XX
SQ Sequence 512 AA;

Query Match		100.0%;	Score 2589;	DB 7;	Length 512;
Best Local Similarity		100.0%;	Pred. No. 5.2e-233;		
Matches 512;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MPAAGSTQORREMAAASAAISGAGRCRLSKIGATRRPPPARVAVRLRPFVDGTAG	60		
DB	1	MPAAGSTQORREMAAASAAISGAGRCRLSKIGATRRPPPARVAVRLRPFVDGTAG	60		
QY	61	ASDPPCVRGMDSCSLEIANRNHQETLKQYDAFYGERSTQODIYAGSVQPIRLHLEQ	120		
DB	61	ASDPPCVRGMDSCSLEIANRNHQETLKQYDAFYGERSTQODIYAGSVQPIRLHLEQ	120		
QY	121	NASVLAYGPTGAGKTHMTLGSPPQGVIPRALMDLLQLTRREGAQRWALSVTMSYLEI	180		
DB	121	NASVLAYGPTGAGKTHMTLGSPPQGVIPRALMDLLQLTRREGAQRWALSVTMSYLEI	180		
QY	181	YQEKVLDLLDPASGDLVIREDCRGNILIPGLSKPTSSPADFERHPLPASRNRVTGATRL	240		
DB	181	YQEKVLDLLDPASGDLVIREDCRGNILIPGLSKPTSSPADFERHPLPASRNRVTGATRL	240		
QY	241	NQSSRSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGSDNRRTGNKGLRKESGAIN	300		
DB	241	NQSSRSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGSDNRRTGNKGLRKESGAIN	300		
QY	301	SLFVLGKVDALNQGLPRVYRDSKLTLLQDSLGSASHTANTAPERFVLTVSAL	360		
DB	301	SLFVLGKVDALNQGLPRVYRDSKLTLLQDSLGSASHTANTAPERFVLTVSAL	360		
QY	361	NFAARKEVINRPTNESLQPHALGPVKLSQKELLGPPEAKRARGPEEEIGSPPEMAAP	420		
DB	361	NFAARKEVINRPTNESLQPHALGPVKLSQKELLGPPEAKRARGPEEEIGSPPEMAAP	420		
QY	421	ASASQKLSPLQKSSMDPAMLERLLSLDRLLASQSGCAPLLSTPKRERVMKTVEEKD	480		
DB	421	ASASQKLSPLQKSSMDPAMLERLLSLDRLLASQSGCAPLLSTPKRERVMKTVEEKD	480		
QY	481	LEIERLTKQKELEAKMLAKAEKEKNCPTM	512		
DB	481	LEIERLTKQKELEAKMLAKAEKEKNCPTM	512		

RESULT 2
ADQ60230
ID ADQ60230 standard; protein; 512 AA.
XX
XX AC ADQ60230;
XX
XX DT 23-SEP-2004 (first entry)
XX
XX DE Human microtubule motor protein #2.
XX
XX KW Human; microtubule motor protein; cellular proliferation disorder;
KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;
KW inflammation; kinsin-like DNA binding protein; KID; autoimmune disease;
KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.
XX
XX OS Homo sapiens.
XX
XX PN US6762043-B1.
XX
XX PD 13-JUL-2004.

XX
PF 06-MAR-2002; 2002US-00093317.
XX
XX 20-APR-1999; 99US-00295612.
PR 20-JUN-2000; 2000US-00597292.
PR 28-NOV-2000; 2000US-00724224.
XX
XX (CYTO-) CYTOKINETICS INC.
PA
XX Beraud C;
XX
XX WPI; 2004-532491/51.
XX
XX New isolated microtubule motor protein, useful for screening modulators
PT for treating cellular proliferation disorders such as cancer,
PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
inflammation.
PT
PS Claim 1; SEQ ID NO 4; 26pp; English.
XX
XX The invention relates to human microtubule motor proteins and the nucleic
CC acids encoding them. The invention also relates to a method of screening
CC for modulators of a motor protein which has microtubule stimulated ATPase
CC activity, a method of testing for ATPase activity of microtubule motor
CC proteins, methods to identify candidate agents that bind to a target
CC protein or act as a modulator of the binding characteristics or
CC biological activity of a target protein, modulators of the target
CC protein, and methods of treating cellular proliferation disorders such as
CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
CC and inflammation, for treating disorders associated with kinsin-like DNA
CC binding protein (KID) and for inhibiting KID. The sequences are used for
CC screening for modulators of motor proteins useful for treating cellular
CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
CC hypertrophy, immune disorders and inflammation, for treating disorders
CC associated with KID and for inhibiting KID and for treating autoimmune
CC diseases, arthritis, graft rejection, inflammatory bowel disease and
CC proliferation induced after medical procedures including surgery and
CC angioplasty. This sequence represents a human microtubule motor protein
CC of the invention. Note: The specification states that this sequence is
CC encoded by the nucleic acid featured as SEQ ID NO:3, but this does not
CC appear to be the case.
XX
XX SQ Sequence 512 AA;

Query Match		100.0%;	Score 2589;	DB 8;	Length 512;
Best Local Similarity		100.0%;	Pred. No. 5.2e-233;		
Matches 512;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MPAAGSTQORREMAAASAAISGAGRCRLSKIGATRRPPPARVAVRLRPFVDGTAG	60		
DB	1	MPAAGSTQORREMAAASAAISGAGRCRLSKIGATRRPPPARVAVRLRPFVDGTAG	60		
QY	61	ASDPPCVRGMDSCSLEIANRNHQETLKQYDAFYGERSTQODIYAGSVQPIRLHLEQ	120		
DB	61	ASDPPCVRGMDSCSLEIANRNHQETLKQYDAFYGERSTQODIYAGSVQPIRLHLEQ	120		
QY	121	NASVLAYGPTGAGKTHMTLGSPPQGVIPRALMDLLQLTRREGAQRWALSVTMSYLEI	180		
DB	121	NASVLAYGPTGAGKTHMTLGSPPQGVIPRALMDLLQLTRREGAQRWALSVTMSYLEI	180		
QY	181	YQEKVLDLLDPASGDLVIREDCRGNILIPGLSKPTSSPADFERHPLPASRNRVTGATRL	240		
DB	181	YQEKVLDLLDPASGDLVIREDCRGNILIPGLSKPTSSPADFERHPLPASRNRVTGATRL	240		
QY	241	NQSSRSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGSDNRRTGNKGLRKESGAIN	300		
DB	241	NQSSRSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGSDNRRTGNKGLRKESGAIN	300		
QY	301	SLFVLGKVDALNQGLPRVYRDSKLTLLQDSLGSASHTANTAPERFVLTVSAL	360		
DB	301	SLFVLGKVDALNQGLPRVYRDSKLTLLQDSLGSASHTANTAPERFVLTVSAL	360		
QY	361	NFAARKEVINRPTNESLQPHALGPVKLSQKELLGPPEAKRARGPEEEIGSPPEMAAP	420		

Db 361 NFAARKEVINPFTNESLQPHALGPVKLSQKELLGPPPEAKRGPEEBEIGSPPEMAAP 420
Qy 421 ASASQKLSPLQKSSMDPAMLERLLSLDRLLASQSGAPLLSTPKRERWLMKTVEEKD 480
Db 421 ASASQKLSPLQKSSMDPAMLERLLSLDRLLASQSGAPLLSTPKRERWLMKTVEEKD 480
Qy 481 LEIERLTKQKELEAKMLAQKAEKENHCPTM 512
Db 481 LEIERLTKQKELEAKMLAQKAEKENHCPTM 512

RESULT 3
ID ADQ09240
AC ADQ09240 standard; protein; 665 AA.
XX ADQ09240;
XX
XX 23-SEP-2004 (first entry)
XX Human KNSL4 protein SEQ ID NO:425.
XX
XX thanatos-associated protein; THAP; THAP responsive gene; THAP family;
XX THAP responsive element; angiogenesis; inflammation; metastasis; cancer;
XX apoptosis; cardiovascular disease; neurodegenerative disease; chemokine;
XX antiangiogenic; antiinflammatory; cardiovascular; cytostatic;
XX neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;
XX human.
XX
XX Homo sapiens.
XX WO2004055050-A2.
XX
XX 01-JUL-2004.
XX
XX 10-DEC-2003; 2003WO-IB006434.
XX
XX 10-DEC-2002; 2002US-0432699P.
XX 03-JUL-2003; 2003US-0485027P.
XX
XX (ENDO-) ENDOCUBE SAS.
XX (CNRS) CNRS CENT NAT RECH SCI.
XX
XX Girard J, Amalric F, Roussigne M, Clouaire T;
XX
XX WPI: 2004-525034/50.
XX N-PSDB; ADQ09241.
XX
XX Modulating expression of a Thanatos (death)-Associated Protein (THAP)
XX responsive gene for preventing or treating e.g. cancer or inflammation,
XX comprises modulating the interaction of a THAP polypeptide with a nucleic
XX acid.
XX
XX Example 47; SEQ ID NO 425; 612pp; English.

The present invention describes a method for modulating the expression of
a thanatos (death)-associated protein (THAP) responsive gene. The method
comprises modulating the interaction of a THAP-family polypeptide or its
biological fragment with a nucleic acid, and so enhancing or repressing
the expression of the THAP responsive gene. Also described: (1) a method
of modulating the expression of a gene responsive to a THAP/chemokine
complex; (2) a pharmaceutical composition comprising a THAP responsive
element in a pharmaceutical carrier; (3) a transcription factor decoy
consisting essentially of a THAP responsive element; (4) a cell
comprising a transcription factor decoy described above; (5) methods of
modulating the interaction between a nucleic acid and a THAP-family
polypeptide or its biological fragment, or a nucleic acid and a
THAP/chemokine complex; (6) a vector packaging cell line comprising a
cell comprising a viral vector which comprises a promoter operably linked
to a nucleic acid encoding a THAP-family polypeptide or its biological
fragment; (7) a method of constructing a cell which expresses a
recombinant THAP-family polypeptide; (8) a method of ameliorating
CC symptoms associated with a condition mediated by a THAP/chemokine complex

CC ; (9) methods of identifying a test compound that modulates transcription
CC at a THAP responsive element or that modulates the transport of a
CC chemokine into the nucleus; (10) methods for reducing the symptoms
CC associated with a condition selected from excessive or insufficient
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
CC or insufficient apoptosis, cardiovascular disease and neurodegenerative
CC diseases; symptoms associated with a condition resulting from the
CC activity of a chemokine or a THAP-family polypeptide in an individual; or
CC symptoms associated with transcriptional repression or activation
CC mediated by a THAP-family polypeptide in an individual; (11) a vector
CC comprising a THAP responsive promoter operably linked to a nucleic acid
CC encoding a detectable product; (12) a genetically engineered cell
CC comprising the vector described above or that expresses a THAP-family
CC polypeptide or its biological fragment; (13) an in vitro transcription
CC reaction comprising a nucleic acid comprising a THAP responsive promoter,
CC ribonucleotides and an RNA polymerase; and (14) an isolated mutant THAP-
CC family polypeptide that does not bind to a chemokine. The pharmaceutical
CC composition has antiangiogenic, antiinflammatory, cardiovascular, excessive
CC cytostatic, neuroprotective and osteopathic activities, and can be used
CC as a THAP and THAP synthesis modulator. The composition can be used for
CC modulating the expression of a THAP responsive gene. Modulation is useful
CC for reducing symptoms of conditions such as excessive or insufficient
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
CC or insufficient apoptosis, cardiovascular disease or neurodegenerative
CC diseases. The present sequence is used in the exemplification of the
XX present invention.

XX Sequence 665 AA;

Query Match 99.5%; Score 2577; DB 8; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.1e-231;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3 AAGSTQORREMAAAASAAISGAGRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGAS 62
Db 2 AAGSTQORREMAAAASAAISGAGRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGAS 61
Qy 63 DPCVRGWDSCSLEIANWRNHQETLKYQDFAFYGERSTQODIYAGSVQPIRLHLEGGNA 122
Db 62 DPCVRGWDSCSLEIANWRNHQETLKYQDFAFYGERSTQODIYAGSVQPIRLHLEGGNA 121
Qy 123 SVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREAGRGPWALSVTMSYLEIYQ 182
Db 122 SVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREAGRGPWALSVTMSYLEIYQ 181
Qy 183 EKVLDDLPASGDLVIREDCRGNLIPGLSQKPISSFADFERHFLPASRNRVTGATRLNQ 242
Db 182 EKVLDDLPASGDLVIREDCRGNLIPGLSQKPISSFADFERHFLPASRNRVTGATRLNQ 241
Qy 243 RSSSHAVLLVKVDORERLAPFRQEGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSI 302
Db 242 RSSSHAVLLVKVDORERLAPFRQEGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSI 301
Qy 303 FVLGKVVVDALNQGLPRVPRYRDSKLTLLQLDSLGSAAHSILIANIAPERRFYLDTSALNF 362
Db 302 FVLGKVVVDALNQGLPRVPRYRDSKLTLLQLDSLGSAAHSILIANIAPERRFYLDTSALNF 361
Qy 363 AARSKEVINPFTNESIQPHALGPVKLSQKELLGPPPEAKRGPEEBEIGSPPEMAAPAS 422
Db 362 AARSKEVINPFTNESIQPHALGPVKLSQKELLGPPPEAKRGPEEBEIGSPPEMAAPAS 421
Qy 423 ASQKLSPLQKSSMDPAMLERLLSLDRLLASQSGAPLLSTPKRERWLMKTVEEKDLE 482
Db 422 ASQKLSPLQKSSMDPAMLERLLSLDRLLASQSGAPLLSTPKRERWLMKTVEEKDLE 481
Qy 483 IERLTKQKELEAKMLAQKAEKENHCPTM 512
Db 482 IERLTKQKELEAKMLAQKAEKENHCPTM 511

RESULT 4
ADQ23344
ID ADQ23344 standard; protein; 487 AA.

XX AC ADC23344;
XX DT 18-DEC-2003 (first entry)
XX DE Human kinesin-like DNA binding protein (KID) (SeqID 8).
XX KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID;
XX KW cytoskeletal; cardiant; immunomodulator; antiinflammatory; gene therapy;
XX KW cancer; hyperplasia; restenosis; cellular proliferation disorder;
XX KW cardiac hypertrophy; immune disorder; inflammation.
XX OS Homo sapiens.
XX PN US6397644-B1.
XX PD 14-MAY-2002.
XX PF 28-NOV-2000; 2000US-00724224.
XX PR 20-APR-1999; 99US-00295612.
XX PR 20-JUN-2000; 2000US-00597292.
XX PA (CYTO-) CYTOKINETICS INC.
XX PI Beraud C;
XX DR WPI; 2003-706919/67.
XX DR N-PSDB; ADC23343.
XX PT Identifying a candidate agent as modulator of function of a target
XX PT protein for treating cellular proliferation disorders by adding a
XX PT candidate agent to a mixture of the target protein that
XX PT directly/indirectly produces ADP or phosphate.
XX PS Claim 1; SEQ ID NO 8; 26pp; English.
XX PS This invention relates to a novel method for high throughput screening
XX CC systems used to identify compounds for the treatment of cellular
XX CC proliferation disorders. Specifically, it refers to candidate agents that
XX CC are capable of modulating the activity of target proteins having motor
XX CC domains, such that the target protein directly or indirectly produces ADP
XX CC or phosphate. Furthermore, this activity can be determined using
XX CC fluorescence or absorbance readouts. The present invention describes a
XX CC method that identifies modulators of the target protein, which is a
XX CC kinesin-like DNA binding protein (known as KID) as cytostatic, cardants,
XX CC immunomodulators and antiinflammatory. Accordingly, through gene
XX CC therapy, they can be used for the treatment of cancer, hyperplasias,
XX CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This
XX CC polypeptide sequence is human KID protein (SeqID 8) of the invention.
XX SQ Sequence 487-AA;

Query Match 95.3%; Score 2467; DB 7; Length 487;
Best Local Similarity 100.0%; Pred. No. 1.3e-221;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 GRCRLSKIGATRRPPPARVAVRLRPFVDGTAGADPPCVRGMDSCSLEIANWRHQT 86
DB 2 GRCRLSKIGATRRPPPARVAVRLRPFVDGTAGADPPCVRGMDSCSLEIANWRHQT 61
QY 87 LKYPQDAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAYGPTGAGKTHMLGSPEQPG 146
DB 62 LKYPQDAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAYGPTGAGKTHMLGSPEQPG 121
QY 147 VIPRALMDLLQLTREBGAERPWALSVTWSYLFYIOEKVLDLDDPASGDLVIREDCRGN 206
DB 122 VIPRALMDLLQLTREBGAERPWALSVTWSYLFYIOEKVLDLDDPASGDLVIREDCRGN 181
QY 207 LIPGLSKPTSSPADFERHFLPASRNRVTGATRLNQRSSRSRAVLVKKVDQERLAPFQ 266
DB 182 LIPGLSKPTSSPADFERHFLPASRNRVTGATRLNQRSSRSRAVLVKKVDQERLAPFQ 241

QY 267 REGKLYLIDLAGSEDNRRITGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVPYRDSKL 326
DB 242 REGKLYLIDLAGSEDNRRITGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVPYRDSKL 301
QY 327 TRLQDSLGSGSAHSILIANIAPERRFYLDTSALNFAARKEVINRPFNTNESIQPHALGP 386
DB 302 TRLQDSLGSGSAHSILIANIAPERRFYLDTSALNFAARKEVINRPFNTNESIQPHALGP 361
QY 387 VKLSQKELLGPPPEAKRARGPEEEEGSPPEMAAPASOKLSPLQKLSMMDPAMLERLLS 446
DB 362 VKLSQKELLGPPPEAKRARGPEEEEGSPPEMAAPASOKLSPLQKLSMMDPAMLERLLS 421
QY 447 LDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDLEIERLTKQKLEAKMLAQAEKE 506
DB 422 LDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDLEIERLTKQKLEAKMLAQAEKE 481
QY 507 NHCPTM 512
DB 482 NHCPTM 487

RESULT 5
ADQ60234
ID ADQ60234 standard; protein; 487 AA.
XX AC ADQ60234;
XX DT 23-SEP-2004 (first entry)
XX DE Human microtubule motor protein #4.
XX KW Human; microtubule motor protein; cellular proliferation disorder;
XX KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;
XX KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;
XX KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.
XX OS Homo sapiens.
XX PN US6762043-B1.
XX PD 13-JUL-2004.
XX PF 06-MAR-2002; 2002US-00093317.
XX PR 20-APR-1999; 99US-00295612.
XX PR 20-JUN-2000; 2000US-00597292.
XX PR 28-NOV-2000; 2000US-00724224.
XX PA (CYTO-) CYTOKINETICS INC.
XX PI Beraud C;
XX DR WPI; 2004-532491/51.
XX DR N-PSDB; ADQ60233.
XX PT New isolated microtubule motor protein, useful for screening modulators
XX PT for treating cellular proliferation disorders such as cancer,
XX PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
XX PT inflammation.
XX PS Claim 1; SEQ ID NO 8; 26pp; English.
XX CC The invention relates to human microtubule motor proteins and the nucleic
XX CC acids encoding them. The invention also relates to a method of screening
XX CC for modulators of a motor protein which has microtubule stimulated ATPase
XX CC activity, a method of testing for ATPase activity of microtubule motor
XX CC proteins, methods to identify candidate agents that bind to a target
XX CC protein or act as a modulator of the binding characteristics or
XX CC biological activity of a target protein, modulators of the target
XX CC protein, and methods of treating cellular proliferation disorders such as
XX CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
XX CC and inflammation, for treating disorders associated with kinesin-like DNA
XX CC binding protein (KID) and for inhibiting KID. The sequences are used for

CC screening for modulators of motor proteins useful for treating cellular
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
 CC hypertrophy, immune disorders and inflammation, for treating disorders
 CC associated with KID and for inhibiting KID and for treating autoimmune
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and
 CC proliferation induced after medical procedures including surgery and
 CC angioplasty. This sequence represents a human microtubule motor protein
 CC of the invention.
 XX
 SQ Sequence 487 AA;

Query Match 95.3%; Score 2467; DB 8; Length 487;
 Best Local Similarity 100.0%; Pred. No. 1.3e-221;
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 GRCLSKIGATRRPPPARVAVLRPPFVDTAGASDPPCVRGMDSCSLEIANWRNHQET 86
 DB 2 GRCLSKIGATRRPPPARVAVLRPPFVDTAGASDPPCVRGMDSCSLEIANWRNHQET 61
 QY 87 LKYPDAFYGERSTQDDIYAGSVQPIRLHLLGGQNASVLAYGPTGAGKTHMLGSPEQFG 146
 DB 62 LKYPDAFYGERSTQDDIYAGSVQPIRLHLLGGQNASVLAYGPTGAGKTHMLGSPEQFG 121
 QY 147 VIPRALMDLLOLTREAGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRNI 206
 DB 122 VIPRALMDLLOLTREAGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRNI 181
 QY 207 LIPGLSQKPISSFADFERHFIAPSNRTVGTALNRSSRSHAVLLVKVDORERLAPRQ 266
 DB 182 LIPGLSQKPISSFADFERHFIAPSNRTVGTALNRSSRSHAVLLVKVDORERLAPRQ 241
 QY 267 REGKLYLDLAGSDNRRTGNKGLRKESGAINTSFLVLGKVDALNQLPRVPYRDSKL 326
 DB 242 REGKLYLDLAGSDNRRTGNKGLRKESGAINTSFLVLGKVDALNQLPRVPYRDSKL 301
 QY 327 TRLLQDSLGGSAHSILIANIAPRRFYLDTSALNFAARSKEVINRPTNESLQPHALGP 386
 DB 302 TRLLQDSLGGSAHSILIANIAPRRFYLDTSALNFAARSKEVINRPTNESLQPHALGP 361
 QY 387 VKLSQKELGPPKARGPPEETGSPPEPAAASQKLSPLQKSSMDPAMLERLLS 446
 DB 362 VKLSQKELGPPKARGPPEETGSPPEPAAASQKLSPLQKSSMDPAMLERLLS 421
 QY 447 LDRLASQSGQAPILSTPKERVMVMTVEKOLIEIRLTKOKELAKMLAKAEKE 506
 DB 422 LDRLASQSGQAPILSTPKERVMVMTVEKOLIEIRLTKOKELAKMLAKAEKE 481
 QY 507 NHCPTM 512
 DB 482 NHCPTM 487

RESULT 6
 ADK40973
 ID ADK40973 standard; protein; 490 AA.
 XX AC
 XX ADK40973;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Novel human kinase protein #80.
 XX
 KW cytosolic; immunomodulator; cardiac; neuroprotective; nootropic;
 KW antiparkinsonian; virucide; antibacterial; fungicide; ophthalmological;
 KW analgesic; hypotensive; immunosuppressive; kinase inhibitor; kinase;
 KW cancer; peripheral nervous system; central nervous system;
 KW Alzheimer's disease; Parkinson's disease; multiple sclerosis;
 KW amyotrophic lateral sclerosis; viral infection; prion infection;
 KW ocular disease; migraine; pain; sexual dysfunction; mood disorder;
 KW attention disorder; cognitive disorder; hypotension; hypertension;
 KW psychotic disorder; neurological disorder; dyskinesia;
 KW metabolic disorder; organ transplant rejection; enzyme.
 XX

OS Homo sapiens.
 XX WO2003057841-A2.
 XX
 XX 17-JUL-2003.
 XX
 XX 31-DEC-2002; 2002WO-US041687.
 XX
 XX 31-DEC-2001; 2001US-0343169P.
 XX
 XX (GRIG/) GRIGORIEV I V.
 XX (SUDA/) SUDARSANAM S.
 XX
 XX Grigoriev IV, Sudarsanam S;
 XX
 XX WPI; 2003-587115/55.
 XX
 XX New isolated, enriched or purified nucleic acid molecule encoding a
 XX kinase polypeptide, useful for treating cancer, immune-related diseases,
 XX cardiovascular disease, brain or neuronal-associated diseases and
 XX metabolic disorders.
 XX
 XX Claim 1; SEQ ID NO 80; 491pp; English.

The invention relates to novel isolated, enriched or purified nucleic acid
 molecules encoding a kinase polypeptide. The nucleic acid molecule
 comprises a sequence that: (a) encodes a kinase polypeptide; (b) is a
 complement of (a); (c) hybridizes under stringent conditions to (a) and
 encodes a naturally occurring kinase polypeptide; (d) encodes the
 polypeptide in (a), except that it lacks one or more, but not all, of an
 N-terminal domain, C-terminal catalytic domain, a catalytic region, a C-
 terminal tail; or (e) is a complement of (d). The nucleic acid molecules,
 polypeptides, methods and substance are useful for treating cancer,
 immune-related diseases or disorders, cardiovascular disease, brain or
 neuronal-associated diseases, and metabolic disorders. The disorders are
 preferably cancers of the tissues or of hematopoietic origin, diseases of
 the central or peripheral nervous system, Alzheimer's disease,
 Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis,
 viral infections, infections caused by prions, infections caused by
 bacteria, infections caused by fungi, ocular diseases, migraines, pain,
 sexual dysfunction, mood disorders, attention disorders, cognition
 disorders, hypotension, hypertension, psychotic disorders, neurological
 disorders, dyskinesias, metabolic disorders and organ transplant
 rejection. This sequence corresponds to one of the kinase polypeptides of
 the invention.

Sequence 490 AA;

Query Match 93.4%; Score 2419; DB 7; Length 490;
 Best Local Similarity 100.0%; Pred. No. 4.1e-217;
 Matches 479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 AAGSTQORREMAAASAAISGAGRCRLSKIGATRRPPPARVAVLRPPFVDTAGAS 62
 DB 10 AAGSTQORREMAAASAAISGAGRCRLSKIGATRRPPPARVAVLRPPFVDTAGAS 69
 QY 63 DPPCVRGMDSCSLEIANWRNHQETLKYQDFAFYGERSTQDDIYAGSVQPIRLHLLGGQNA 122
 DB 70 DPPCVRGMDSCSLEIANWRNHQETLKYQDFAFYGERSTQDDIYAGSVQPIRLHLLGGQNA 129
 QY 123 SVLAYGPTGAGKTHMLGSPEQGVIPRALMDLLOLTREAGAGRPWALSVTMSYLEIYQ 182
 DB 130 SVLAYGPTGAGKTHMLGSPEQGVIPRALMDLLOLTREAGAGRPWALSVTMSYLEIYQ 189
 QY 183 EKVLDDLPASGDLVIREDCRGNILIPGLSQKPISSFADFERHFIAPSNRTVGTALN 242
 DB 190 EKVLDDLPASGDLVIREDCRGNILIPGLSQKPISSFADFERHFIAPSNRTVGTALN 249
 QY 243 RSSRSHAVLLVKVDORERLAPFRQEGKLYLDLAGSDNRRTGNKGLRKESGAINTS 302
 DB 250 RSSRSHAVLLVKVDORERLAPFRQEGKLYLDLAGSDNRRTGNKGLRKESGAINTS 309

QY 303 FVLGVVDALNOGLPRVPRYRDSKLTLLQDSLGSSAHSILIANIAPERRFYLDTVSALNF 362
DB 310 FVLGVVDALNOGLPRVPRYRDSKLTLLQDSLGSSAHSILIANIAPERRFYLDTVSALNF 369
QY 363 AARSKEVINPFTNESIQPHALGPVKLSQKELGPPPEAKRARGPEEEIGSPPEMAAPAS 422
DB 370 AARSKEVINPFTNESIQPHALGPVKLSQKELGPPPEAKRARGPEEEIGSPPEMAAPAS 429
QY 423 ASQKLSPLQKLSMDPAMLERLLSLDRLLASQSQGAPLLSTPKRERMVLMKTVEEKOL 481
DB 430 ASQKLSPLQKLSMDPAMLERLLSLDRLLASQSQGAPLLSTPKRERMVLMKTVEEKOL 488

RESULT 7
AAB56650
ID AAB56650 standard; protein; 460 AA.
XX AAB56650;
AC AAB56650;
XX 13-MAR-2001 (first entry)
XX Human prostate cancer antigen protein sequence SEQ ID NO:1228.
DE Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
XX neuroprotective; cytostatic; cardioprotective; immunomodulatory; muscular;
XX vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
XX antibacterial; gene therapy; neural; immune; reproductive; renal;
XX gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX wound; infectious disease.
XX Homo sapiens.
OS Homo sapiens.
XX WO200055174-A1.
XX 21-SEP-2000.
XX 08-MAR-2000; 2000WO-US005988.
XX 12-MAR-1999; 99US-0124270P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX Rosen CA, Ruben SM;
FI WPI; 2000-587513/55.
XX N-PSDB; AAF15853.
XX Prostate cancer associated gene sequences, referred to as prostate cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as prostate cancer.
XX Claim 11; Page 1649-1651; 2338pp; English.
XX AAF15566 to AAF16505 encode the human prostate cancer associated
XX proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
XX The prostate cancer antigens can have neuroprotective, cytostatic,
XX cardioprotective, immunomodulatory, muscular, vulnary, gastrointestinal,
XX nephrotropic, antiinfective, gynaecological and antibacterial activities,
XX and can be used in gene therapy. The prostate cancer antigen
XX polynucleotides may be used for detection of prostate cancer, chromosome
XX identification, as chromosome markers, and for numerous other diagnostic
XX or research purposes. The prostate cancer antigens may be used to treat
XX disorders such as neural, immune, muscular, reproductive,
XX gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
XX AAB57303 represent sequences used in the exemplification of the present
XX invention
XX Sequence 460 AA;

Query Match 83.1%; Score 2151; DB 3; Length 460;
Best Local Similarity 95.9%; Pred. No. 4.6e-192;

Matches 424; Conservative 1; Mismatches 17; Indels 0; Gaps 0;
QY 25 GAGRCRLSKI GATRRPPPPARVRVRLPPFVDDGTAGASDPPCVRGMDSCSLEIANWRNHQ 84
DB 2 GAGRCRLSKI GATRRPPPPARVRVRLPPFVDDGTAGASDPPCVRGMDSCSLEIANWRNHQ 61
QY 85 ETLKYQDFAFYGERSTQODIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHMLGSPEQ 144
DB 62 ETLKYQDFAFYGERSTQODIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHMLGSPEQ 121
QY 145 PGVIPRALMDLLQLTFRREGAERPMWLSVTWMSYLEIYQEKVLDLLDPASGDLVIREDCRG 204
DB 122 PGVIPRALMDLLQLTFRREGAERPMWLSVTWMSYLEIYQEKVLDLLDPASGDLVIREDCRG 181
QY 205 NILIPGLSQKPISSPADFERHFLPASRNRVTGATRLNQRSSRSRSHAVLLVKVDQERLAPF 264
DB 182 NILIPGLSQKPISSPADFERHFLPASRNRVTGATRLNQRSSRSRSHAVLLVKVDQERLAPF 241
QY 265 ROREGKLYLIDLAGEDNRRRTGNKRLKESGAINTSFLVLGVVDALNOGLPRVPRYRDS 324
DB 242 ROREGKLYLIDLAGEDNRRRTGNKRLKESGAINTSFLVLGVVDALNOGLPRVPRYRDS 301
QY 325 KLTRLQLDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVINPFTNESIQPHAL 384
DB 302 KLTRLQLDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVINPFTNESIQPHAL 361
QY 385 GPVKLSQKELGPPPEAKRARGPEEEIGSPPEMAAPASOKLSPLQKLSMDPAMLERL 444
DB 362 GPVKLSQKELGPPPEAKRARGPEEEIGSPPEMAAPASOKLSPLQKLSMDPAMLERL 421
QY 445 LSLDRLLASQSQGAPLLSTPK 466
DB 422 LQLGPSACLPGEPXGFSVEYPK 443

RESULT 8
ADC23338
ID ADC23338 standard; protein; 370 AA.
XX ADC23338;
AC ADC23338;
XX 18-DEC-2003 (first entry)
XX Human kinesin-like DNA binding protein (KID) (SeqID 2).
DE Human; enzyme; motor domain; kinesin-like DNA binding protein; KID;
XX cytostatic; cardiant; immunomodulator; antiinflammatory; gene therapy;
XX cancer; hyperplasia; restenosis; cellular proliferation disorder;
XX cardiac hypertrophy; immune disorder; inflammation.
XX Homo sapiens.
OS Homo sapiens.
XX Key Location/Qualifiers
FT Misc-difference 1 /note= "Encoded by ATGCA"
FT US6387644-B1.
XX PN US6387644-B1.
XX 14-MAY-2002.
XX PD 14-MAY-2002.
XX PF 28-NOV-2000; 2000US-0072422A.
XX PR 20-APR-1999; 99US-00295612.
XX PR 20-JUN-2000; 2000US-00597292.
XX PA (CYTO-) CYTOKINETICS INC.
XX PI Beraud C;
XX WPI; 2003-706919/67.
XX N-PSDB; ADC23337.
XX Identifying a candidate agent as modulator of function of a target
PT

PT protein for treating cellular proliferation disorders by adding a
PT candidate agent to a mixture of the target protein that
PT directly/indirectly produces ADP or phosphate.

XX PS Claim 1; SEQ ID NO 2; 26pp; English.

XX CC This invention relates to a novel method for high throughput screening
XX systems used to identify compounds for the treatment of cellular
XX proliferation disorders. Specifically, it refers to candidate agents that
XX are capable of modulating the activity of target proteins having motor
XX domains, such that the target protein directly or indirectly produces ADP
XX or phosphate. Furthermore, this activity can be determined using
XX fluorescence or absorbance readouts. The present invention describes a
XX method that identifies modulators of the target protein, which is a
XX kinesin-like DNA binding protein (known as KID) as cytosolic, cardinals,
XX immunomodulators and antiinflammatory. Accordingly, through gene
XX therapy, they can be used for the treatment of cancer, hyperplasias,
XX restenosis, cardiac hypertrophy, immune disorders and inflammation. This
XX polypeptide sequence is human KID protein (SeqID 2) of the invention.

XX SQ Sequence 370 AA;

Query Match 72.3%; Score 1873; DB 7; Length 370;

Best Local Similarity 100.0%; Pred. No. 3.5e-166; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AAGSTQORREMAAASAAISGAGRCRLSKIGATRRPPARVRVAVRLRPFVDGTAGAS 62
Db 2 AAGSTQORREMAAASAAISGAGRCRLSKIGATRRPPARVRVAVRLRPFVDGTAGAS 61
Qy 63 DPPCVRGMDSCSLEIANWRNHQETLKYQDAFYGERSTQDDIYAGSVQPIRLHLLGQNA 122
Db 62 DPPCVRGMDSCSLEIANWRNHQETLKYQDAFYGERSTQDDIYAGSVQPIRLHLLGQNA 121
Qy 123 SVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTREEGAGRPWALSVMYLEIYQ 182
Db 122 SVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTREEGAGRPWALSVMYLEIYQ 181
Qy 183 EKVLDDLPASGDLVIREDCRGNILIPGLSQKPISSFADFRRHFLPASRNTVATRLNQ 242
Db 182 EKVLDDLPASGDLVIREDCRGNILIPGLSQKPISSFADFRRHFLPASRNTVATRLNQ 241
Qy 243 RSSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSL 302
Db 242 RSSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSL 301
Qy 303 FVLGKVVDALNQGLPRVPYRDSKLTLLQDSLGGSAHSILIANIAPERRFYLDTVSALNF 362
Db 302 FVLGKVVDALNQGLPRVPYRDSKLTLLQDSLGGSAHSILIANIAPERRFYLDTVSALNF 361
Qy 363 AARSKEVIN 371
Db 362 AARSKEVIN 370

RESULT 9

ID ADQ60228 standard; protein; 370 AA.

XX AC ADQ60228;

XX DT 23-SEP-2004 (first entry)

XX DE Human microtubule motor protein #1.

KW Human; microtubule motor protein; cellular proliferation disorder;
KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;
KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;
KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.

OS Homo sapiens.

XX US6762043-B1.

PN

XX 13-JUL-2004.

XX PF 06-MAR-2002; 2002US-00093317.

XX PR 20-APR-1999; 99US-00295612.

XX PR 20-JUN-2000; 2000US-00597292.

XX PR 28-NOV-2000; 2000US-00724224.

XX PA (CYTO-) CYTOKINETICS INC.

XX PS Beraud C;

XX WIPI; 2004-532491/51.

XX PT New isolated microtubule motor protein, useful for screening modulators
XX for treating cellular proliferation disorders such as cancer,
XX hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
XX inflammation.

XX PS Claim 1; SEQ ID NO 2; 26pp; English.

XX CC The invention relates to human microtubule motor proteins and the nucleic
XX acids encoding them. The invention also relates to a method of screening
XX for modulators of a motor protein which has microtubule stimulated ATPase
XX activity, a method of testing for ATPase activity of microtubule motor
XX proteins, methods to identify candidate agents that bind to a target
XX protein or act as a modulator of the binding characteristics or
XX biological activity of a target protein, modulators of the target
XX protein, and methods of treating cellular proliferation disorders such as
XX cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
XX and inflammation, for treating disorders associated with kinesin-like DNA
XX binding protein (KID) and for inhibiting KID. The sequences are used for
XX screening for modulators of motor proteins useful for treating cellular
XX proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
XX hypertrophy, immune disorders and inflammation, for treating disorders
XX associated with KID and for inhibiting KID and for treating autoimmune
XX diseases, arthritis, graft rejection, inflammatory bowel disease and
XX angioplasty. This sequence represents a human microtubule motor protein
XX of the invention. Note: The specification states that this sequence is
XX encoded by the nucleic acid featured as SEQ ID NO:1, but this does not
XX appear to be the case.

XX SQ Sequence 370 AA;

Query Match 72.3%; Score 1873; DB 8; Length 370;

Best Local Similarity 100.0%; Pred. No. 3.5e-166;

Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AAGSTQORREMAAASAAISGAGRCRLSKIGATRRPPARVRVAVRLRPFVDGTAGAS 62
Db 2 AAGSTQORREMAAASAAISGAGRCRLSKIGATRRPPARVRVAVRLRPFVDGTAGAS 61
Qy 63 DPPCVRGMDSCSLEIANWRNHQETLKYQDAFYGERSTQDDIYAGSVQPIRLHLLGQNA 122
Db 62 DPPCVRGMDSCSLEIANWRNHQETLKYQDAFYGERSTQDDIYAGSVQPIRLHLLGQNA 121
Qy 123 SVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTREEGAGRPWALSVMYLEIYQ 182
Db 122 SVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTREEGAGRPWALSVMYLEIYQ 181
Qy 183 EKVLDDLPASGDLVIREDCRGNILIPGLSQKPISSFADFRRHFLPASRNTVATRLNQ 242
Db 182 EKVLDDLPASGDLVIREDCRGNILIPGLSQKPISSFADFRRHFLPASRNTVATRLNQ 241
Qy 243 RSSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSL 302
Db 242 RSSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSL 301
Qy 303 FVLGKVVDALNQGLPRVPYRDSKLTLLQDSLGGSAHSILIANIAPERRFYLDTVSALNF 362
Db 302 FVLGKVVDALNQGLPRVPYRDSKLTLLQDSLGGSAHSILIANIAPERRFYLDTVSALNF 361

QY 363 AARSKEVIN 371
Db 362 AARSKEVIN 370

RESULT 10
ADC23342
ID ADC23342 standard; protein; 346 AA.
XX AC ADC23342;
XX DT 18-DEC-2003 (first entry)
XX DE Human kinesin-like DNA binding protein (KID) (SeqID 6).
XX human; enzyme; motor domain; kinesin-like DNA binding protein; KID;
KW cytoskeletal; cardiant; immunomodulator; antiinflammatory; gene therapy;
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;
KW cardiac hypertrophy; immune disorder; inflammation.
XX OS Homo sapiens.
XX FN US6387644-B1.
XX PD 14-MAY-2002.
XX PF 28-NOV-2000; 2000US-00724224.
XX PR 20-APR-1999; 99US-00295612.
XX PR 20-JUN-2000; 2000US-00597292.
XX PA (CYTO-) CYTOKINETICS INC.
XX PI Beraud C;
XX WPI; 2003-706919/67.
XX DR N-PSDB; ADC23341.
XX PT Identifying a candidate agent as modulator of function of a target
XX protein for treating cellular proliferation disorders by adding a
XX candidate agent to a mixture of the target protein that
XX directly/indirectly produces ADP or phosphate.
XX PS Claim 1; SEQ ID NO 6; 26pp; English.
XX This invention relates to a novel method for high throughput screening
XX systems used to identify compounds for the treatment of cellular
XX proliferation disorders. Specifically, it refers to candidate agents that
XX are capable of modulating the activity of target proteins having motor
XX domains, such that the target protein directly or indirectly produces ADP
XX or phosphate. Furthermore, this activity can be determined using
XX fluorescence or absorbance readouts. The present invention describes a
XX method that identifies modulators of the target protein, which is a
XX kinesin-like DNA binding protein (known as KID) as cytostatic, cardiant,
XX immunomodulators and antiinflammatory. Accordingly, through gene
XX therapy, they can be used for the treatment of cancer, hyperplasias,
XX restenosis, cardiac hypertrophy, immune disorders and inflammation. This
XX polypeptide sequence is human KID protein (SeqID 6) of the invention.
SQ Sequence 346 AA;

Query Match 68.1%; Score 1763; DB 7; Length 346;
Best Local Similarity 100.0%; Pred. No. 6.4e-156;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 GCRLSKIGATRRPPPARVRVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQET 86
Db 2 GCRLSKIGATRRPPPARVRVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61

QY 87 LKYQDFAFYGERSTQODIYAGSVQPIRLHLEGQNASVLAVGPTCAGKTHMLGSPQPG 146
Db 62 LKYQDFAFYGERSTQODIYAGSVQPIRLHLEGQNASVLAVGPTCAGKTHMLGSPQPG 121

QY 147 VIPRALMDLLQLTREAGRPNWALSVTMSYLIYQEKVLDLDPASGDLVIREDCGNI 206
Db 122 VIPRALMDLLQLTREAGRPNWALSVTMSYLIYQEKVLDLDPASGDLVIREDCGNI 181

QY 207 LIPGLSOKPISSPADFERHFLPASRNRVTGATRLNQRSSSHAVLLVKVDQERLAPFRQ 266
Db 182 LIPGLSOKPISSPADFERHFLPASRNRVTGATRLNQRSSSHAVLLVKVDQERLAPFRQ 241

QY 267 REGKLYLIDLAGSEDDNRRTGNKGLRLKESGAINTSFLVGLKVVQDALNQGLPRPYRDSKL 326
Db 242 REGKLYLIDLAGSEDDNRRTGNKGLRLKESGAINTSFLVGLKVVQDALNQGLPRPYRDSKL 301

QY 327 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 371
Db 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

RESULT 11
ADQ60232
ID ADQ60232 standard; protein; 346 AA.
XX AC ADQ60232;
XX DT 23-SEP-2004 (first entry)
XX DE Human microtubule motor protein #3.
XX Human; microtubule motor protein; cellular proliferation disorder;
KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;
KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;
KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.
XX OS Homo sapiens.
XX FN US6762043-B1.
XX PD 13-JUL-2004.
XX PF 06-MAR-2002; 2002US-00093317.
XX PR 20-APR-1999; 99US-00295612.
XX PR 20-JUN-2000; 2000US-00597292.
XX PR 28-NOV-2000; 2000US-00724224.
XX PA (CYTO-) CYTOKINETICS INC.
XX PI Beraud C;
XX WPI; 2004-532491/51.
XX DR N-PSDB; ADQ60231.
XX PT New isolated microtubule motor protein, useful for screening modulators
XX for treating cellular proliferation disorders such as cancer,
XX hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
XX inflammation.
XX PS Claim 1; SEQ ID NO 6; 26pp; English.
XX The invention relates to human microtubule motor proteins and the nucleic
XX acids encoding them. The invention also relates to a method of screening
XX for modulators of a motor protein which has microtubule stimulated ATPase
XX activity, a method of testing for ATPase activity of microtubule motor
XX proteins, methods to identify candidate agents that bind to a target
XX protein or act as a modulator of the binding characteristics or
XX biological activity of a target protein, modulators of the target
XX protein, and methods of treating cellular proliferation disorders such as
XX cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
XX and inflammation, for treating disorders associated with kinesin-like DNA
XX binding protein (KID) and for inhibiting KID. The sequences are used for
XX screening for modulators of motor proteins useful for treating cellular
XX proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
XX hypertrophy, immune disorders and inflammation, for treating disorders

CC associated with KID and for inhibiting KID and for treating autoimmune
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and
 CC proliferation induced after medical procedures including surgery and
 CC angioplasty. This sequence represents a human microtubule motor protein
 CC of the invention.

XX
 SQ Sequence 346 AA;
 Query Match 68.1%; Score 1763; DB 8; Length 346;
 Best Local Similarity 100.0%; Pred. No. 6.4e-156;
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 27 GRCRLSKIGATRRPPPARVRVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANRNHOET 86
 DB |||||
 QY 87 LKYQFDAYFGERSQODIYAGSVQPIRLHLEGGQNASVLAQPTGAGKTHMLGSPQPG 146
 DB |||||
 QY 62 LKYQFDAYFGERSQODIYAGSVQPIRLHLEGGQNASVLAQPTGAGKTHMLGSPQPG 121
 DB |||||
 QY 147 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 206
 DB |||||
 QY 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 181
 DB |||||
 QY 207 LIPGLSQKPISSFADFERRHFLPASRNRVTGATRLNQRSSRSRAVLLVKVDQRLAPPRQ 266
 DB |||||
 QY 182 LIPGLSQKPISSFADFERRHFLPASRNRVTGATRLNQRSSRSRAVLLVKVDQRLAPPRQ 241
 DB |||||
 QY 267 REGKLYLDLAGSEDNRRTGNKGLKESGAINTSFLVGLKVDALNQLPRVPYRQSKL 326
 DB |||||
 QY 242 REGKLYLDLAGSEDNRRTGNKGLKESGAINTSFLVGLKVDALNQLPRVPYRQSKL 301
 DB |||||
 QY 327 TRLLQDSLGGSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 371
 DB 302 TRLLQDSLGGSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 346

RESULT 12
 ABB71112
 ID ABB71112 standard; protein; 784 AA.
 XX
 AC ABB71112;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 40128.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.
 XX Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL15215.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 40128; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 784 AA;
 Query Match 23.8%; Score 616; DB 4; Length 784;
 Best Local Similarity 33.1%; Pred. No. 6.1e-48;
 Matches 176; Conservative 84; Mismatches 183; Indels 88; Gaps 18;
 QY 37 TRRP-----PPARVRVAVRLRPFDGTAGASDPPCV-----RGMDSCSLEIANW--RN 82
 DB |||||
 QY 5 SRRFGTSSQTPNECVVVRCPMSNRERSERSPEVVVVPNRGV-----VELQNVVDGN 60
 DB |||||
 QY 83 HQETLKVFQDFYGERSTQODIYAGSVQPIRLHLEGGQNASVLAQPTGAGKTHM--L 139
 DB |||||
 QY 61 KEQRKVFTYDAAYDASATQTLTLYHEVFPVSVLEGFNGCIFYAGTGTGKTFTMEGVR 120
 DB |||||
 QY 140 GSPEQPGVIRALMDL-LQLTREAGRGPWALSVTMSYLEIYQEKVLDLDPASGDLVI 198
 DB |||||
 QY 121 GNDELGMGIIPTFEQIWLHINRTN-----FQFLVDVSYLIEYMEELRDLKPNKSHLEV 175
 DB |||||
 QY 199 REDCRGN-ILIPGLSQKPISSFADFERRHFLPASRNRVTGATRLNQRSSRSRAVLLVKVDQ 257
 DB |||||
 QY 176 RE--RGSGVVPVNLHAINCKSVEDMIKVMQVGNKRTVGTVMNEHSSRSRAIFMIKEM 233
 DB |||||
 QY 258 RERLAPPRQREGKLYLDLAGSEDNRRTGNKGLKESGAINTSFLVGLKVDALNQLP 317
 DB |||||
 QY 234 CD-TETNTIKVGNLIDLAGSERQSKTGAERLKEASKINLALSLGNVISALAESPP 292
 DB |||||
 QY 318 RVPYRDSKLTLLQDSLGGSAHSILIANIAPERFYLDTVSALNFAARSKEVINPPTNE 377
 DB |||||
 QY 293 HVPYRDSKLTLLQDSLGGSKTMIANIGSNYNYNETLTTLRYAGRAKSIQNPQKNE 352
 DB |||||
 QY 378 SLOPHALGPVKLSQ-----KELLGPPAKRA-----RGPEEEI----- 411
 DB |||||
 QY 353 DPQ-----DAKLKEYQEEIERLKLIGPQQQORSEKQVAKKQVKKPKETVTKEMSDS 407
 DB |||||
 QY 412 -----GSPFMAAPASAKLSPLQKLSMDPAMLERLSLDRLLASQGSQG 458
 DB |||||
 QY 408 LQVSTIEQPVDDSDPE---GAESSEKENEAEVAKSNEELERERVEN-SKLAALAELE 463
 DB |||||
 QY 459 APLLSTPKRERMVLMKTVEEKLIEIER-----LTKQKELEAKMLAQKAE 504
 DB |||||
 QY 464 GQLVRGKN-----LLDTYSEKQIELEKLVIAERKKREIEIQOOLEQEE 510
 DB |||||
 RESULT 13
 ABB65183
 ID ABB65183 standard; protein; 677 AA.
 XX
 AC ABB65183;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 22341.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.
 XX Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX N-PSDB; ABL09286.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.
XX Disclosure; SEQ ID NO 22341; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 677 AA;
Query Match 22.9%; Score 593; DB 4; Length 677;
Best Local Similarity 33.9%; Pred. No. 6.9e-46;
Matches 167; Conservative 78; Mismatches 190; Indels 58; Gaps 12;
QY 45 VRVAVRLRPFDG--TAGADPPCVRGMD--SCSLEIANWNRHQETLKYQFDAPYGERSTQ 101
DB 21 VRVVRTRPMDKNELSAGALSIVDKINRAITVMKPNATANPPTTYTFDNNVFDGGSNQ 80
QY 102 QDIYAGSVQPIRLHLEGGNASVLAYGPTGAGTHTMLGSPEOP---GVIPRALMDL--- 155
DB 81 MDLYVDTRIVDKVLENGYNGTILAYGQTGTGTYTMSGNPDSPQTKGIIPNAFAHIFGH 140
QY 156 LQITREGEAGRPWALSVTWMSYLEIYOEKVLDLDPASG--DLVIREDCRGNILIPGLSOK 214
DB 141 IAKAKEN-----QKFLVRVSYMEIYNEEVRDLGKDVGKSLVKERPDIGVFVKDLSGY 194
QY 215 PISFADFERHFLPASNRRTVGATRLNQRSSRHAVLLVKVDORERLAPRQ--REGKLY 272
DB 195 VVHNADDLENIMRLGNKRVAGATGMNQSRSRSHAFITVSELSGEGDVQVVRMGKLQ 254
QY 273 LIDLAGSDNRRTGNKGLRKESGAINTSFLVGLGVVDALNQG--LPRVPVYRDSKLTRELQ 331
DB 255 LVDLAGSEROKTQASQRLKEATKINLSVLGNVISALVDGKSTHIPPYNSKLTRELQ 314
QY 332 DSLGGSASHILIANIAPRRFYDVTVALNFAARKEVINRPTNE-----SLOPHALGPV 387
DB 315 DSLGNSKTVCATISPADSNYMETITSLRYASRAKNIQNRMINEEPKDALLRHFQEEI 374
QY 388 KLSOKEL-LGPPEAKRARGPEERIGSPPEWAPASASQKLSPLQKLSMDPAMLERLLS 446
DB 375 ARLRKQLEGDUSLEERPPSEEDTADDELEAPLELETSSTIQAVEKKPKKREK--- 431
QY 447 LDRLLASQGSQAPLLSTPKRRRMVLMKTVEEKDLEIRLTKQKLEAKM----- 497
DB 432 -----TDAEKELAKRNEHQIEHAQTEOTLRNKLVSLEGIKLV 473
QY 498 -----LAQKAEKE 506
DB 474 GGENLEKAQTOE 486

RESULT 14
ADK40926
ID ADK40926 standard; protein; 1237 AA.
XX
AC ADK40926;
XX
DT 06-MAY-2004 (first entry)
XX
DE Novel human kinase protein #33.
XX
KW cytotatic; immunomodulator; cardiant; neuroprotective; nootropic;
KW antiparkinsonian; virucide; antibacterial; fungicide; ophthalmological;
KW analgesic; hypotensive; immunosuppressive; kinase inhibitor; kinase;
KW cancer; peripheral nervous system; central nervous system;
KW Alzheimer's disease; Parkinson's disease; multiple sclerosis;
KW amyotrophic lateral sclerosis; viral infection; prion infection;
KW ocular disease; migraine; pain; sexual dysfunction; mood disorder;
KW attention disorder; cognition disorder; hypotension; hypertension;
KW psychotic disorder; neurological disorder; dyskinesia;
KW metabolic disorder; organ transplant rejection; enzyme.
XX
OS Homo sapiens.
XX
PN WO2003057841-A2.
XX
PD 17-JUL-2003.
XX
PF 31-DEC-2002; 2002WO-US041687.
XX
PR 31-DEC-2001; 2001US-0343169P.
XX
PA (GRIG/) GRIGORIEV I V.
PA (SUDA/) SUDARSANAM S.
PI Grigoriev IV, Sudarsanam S;
XX
XX WPI; 2003-587115/55.
XX
XX New isolated, enriched or purified nucleic acid molecule encoding a
XX kinase polypeptide, useful for treating cancer, immune-related diseases,
XX cardiovascular disease, brain or neuronal-associated diseases and
XX metabolic disorders.
XX
XX Claim 1; SEQ ID NO 33; 491pp; English.
XX
XX The invention relates to novel isolated, enriched or purified nucleic acid
XX molecules encoding a kinase polypeptide. The nucleic acid molecule
XX comprises a sequence that: (a) encodes a kinase polypeptide; (b) is a
XX complement of (a); (c) hybridizes under stringent conditions to (a) and
XX encodes a naturally occurring kinase polypeptide; (d) encodes the
XX polypeptide in (a), except that it lacks one or more, but not all, of an
XX N-terminal domain, C-terminal catalytic domain, a catalytic domain, a C-
XX terminal domain, a coiled-coil structure region, a spacer region and a C-
XX terminal tail; or (e) is a complement of (d). The nucleic acid molecules,
XX polypeptides, methods and substance are useful for treating cancers,
XX immune-related diseases or disorders, cardiovascular disease, brain or
XX neuronal-associated diseases, and metabolic disorders. The disorders are
XX preferably cancers of the tissues or of hematopoietic origin, diseases of
XX the central or peripheral nervous system, Alzheimer's disease,
XX Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis,
XX viral infections, infections caused by prions, infections caused by
XX bacteria, infections caused by fungi, ocular diseases, migraines, pain,
XX sexual dysfunction, mood disorders, attention disorders, cognitive
XX disorders, hypotension, hypertension, psychotic disorders, neurological
XX disorders, dyskinesias, metabolic disorders and organ transplant
XX rejection. This sequence corresponds to one of the kinase polypeptides of
XX the invention.
XX
XX Sequence 1237 AA;
Query Match 22.3%; Score 578.5; DB 7; Length 1237;
Best Local Similarity 32.4%; Pred. No. 4.1e-44;
Matches 157; Conservative 97; Mismatches 172; Indels 59; Gaps 13;

Db 286 VPYRDKLRLQLQDSLGGNSHTLMIACVSPADSNLEETLNTLRYADRAKKIKNKPVIMID 345
Qy 379 LQPHALGPVK--LSQKELLGPPPEAKRGPEEEEEIGSPPEMAAPASASOKLSP-LQKLSS 435
Db 346 PQPAELNHLKQVQQLQVL-----LQAHGG-----TLPGSITVEPSENQSLME 390
Qy 436 MDPAMLERLLSLDRLLASQGSQCAPLLSTPKRFRVLMKTVEEK-DLEIERLKTOK-EL 493
Db 391 KNQSLVEENEKLSRGLSEAGQTAQML-----ERILTEQANEKNNAKLEELRQHAACKL 445
Qy 494 EAKMLAQKAEKE 506
Db 446 DLQKLVELLEDOE 458

Search completed: November 5, 2004, 18:44:39
Job time : 107.087 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 ; Search time 10.7872 Seconds
(without alignments)
3300.235 Million cell updates/sec

Title: US-10-797-893-2

Perfect score: 1878

Sequence: 1 MAAGSTQORREMAAASAA.....FYLDVTVSALNFAARSKEVIN 370

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1665	88.7	665	S62328	kinesin-like DNA b
2	594.5	31.7	664	T48258	kinesin-like prote
3	555.5	29.6	784	A55236	kinesin-related pr
4	550	29.3	703	A86319	F25116.11 protein
5	547	29.1	1254	T18277	kinesin heavy chai
6	536.5	28.6	1058	T47525	kinesin-related pr
7	536	28.5	784	T40594	probable kinesin-1
8	532	28.3	1226	I51617	kinesin-like prote
9	531.5	28.3	885	D86151	F22M8.8 protein -
10	527.5	28.1	1388	T30335	KLP2 protein - Afr
11	525	28.0	1070	T06733	kinesin homolog F2
12	524	27.9	1056	H84777	probable kinesin-r
13	521.5	27.8	1006	T02017	kinesin-related pr
14	517.5	27.6	1076	B84687	probable kinesin-1
15	514.5	27.4	699	S38982	kinesin-related pr
16	514	27.4	1696	A48669	kinesin-related pr
17	514	27.4	1695	A56921	kinesin family pro
18	513.5	27.3	701	B44259	kinesin-related pr
19	513.5	27.3	805	S64238	kinesin-related pr
20	513.5	27.3	1225	A56514	chromokinesin - ch
21	513	27.3	747	A57107	kinesin-related pr
22	511.5	27.2	1022	E84792	probable kinesin h
23	509.5	27.1	786	A53939	kinesin homolog KH
24	509.5	27.1	2954	T14156	kinesin-related pr
25	508	27.1	1229	T48959	kinesin-like prote
26	507.5	27.0	1231	A54803	microtubule-associ
27	507	27.0	928	T10164	kinesin heavy chai
28	504.5	26.9	330	B48835	kinesin-like prote
29	504	26.8	813	T46242	kinesin-like prote

30	501	26.7	1056	1	G02157	kinesin-like apind
31	500	26.6	1150	1	A55289	kinesin-like prote
32	498.5	26.5	935	2	T51930	kinesin [imported]
33	495.5	26.4	958	2	T20621	hypothetical prote
34	495	26.4	929	2	T51932	kinesin [imported]
35	494.5	26.3	883	2	T40128	kinesin-like prote
36	491	26.1	963	1	A41919	kinesin heavy chai
37	490	26.1	742	1	S58691	kinesin-related pr
38	488.5	26.0	1085	2	T38378	kinesin-like prote
39	488	26.0	987	2	T51360	kinesin-like heavy
40	485.5	25.9	793	2	JC5831	kinesin-related pr
41	485.5	25.9	843	2	S44868	kinesin heavy chai
42	483	25.7	332	2	C48835	kinesin-like prote
43	482.5	25.7	1584	1	JN0114	kinesin-related pr
44	482.5	25.7	1584	2	T15822	kinesin-like prote
45	482	25.7	1027	2	S37711	kinesin heavy chai

ALIGNMENTS

RESULT 1

S62328

kinesin-like DNA binding protein KID - human

C:Species: Homo sapiens (man)

C>Date: 23-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 16-Aug-2004

C:Accession: S62328

R:Tokai, N.; Fujimoto-Nishiyama, A.; Toyoshima, Y.; Yonemura, S.; Tsukita, S.; Inoue, J.

EMBO J. 15, 457-467, 1996

A:Title: Kid, a novel kinesin-like DNA binding protein, is localized to chromosomes and

A:Reference number: S62328; MUID:96174806; PMID:8599929

A:Accession: S62328

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-665 <OK>

C:Superfamily: kinesin motor domain homology

C:Keywords: ATP; nucleotide binding; P-loop

F:44-374/Domain: kinesin motor domain homology <KMOT>

F:128-135/Region: nucleotide-binding motif A (P-loop)

Query Match 88.7%; Score 1665; DB 2; Length 665;

Best Local Similarity 91.4%; Pred. No. 2.4e-125;

Matches 339; Conservative 1; Mismatches 29; Indels 2; Gaps 2;

Qy	1	MAAGSTQORREMAAASAAISGRCRLSKIGATRRPPPARVAVVLRFPVDTAGA	60
Db	1	MAAGSTQORREMAAASAAISGRCRLSKIGATRRPPPARVAVVLRFPVDTAGA	60
Qy	61	SDPPCVRGMDSCSLEIANWRNHQETLKQFDAFYGERSTQDDIYAGSVQPIRLHLEGGN	120
Db	61	SDPPCVRGMDSCSLEIANWRNHQETLKQFDAFYGERSTQDDIYAGSVQPIRLHLEGGN	120
Qy	121	AS-VLAYGPTGAGKTHTMLGSPQPGVTPRALMDLLQLTREBAGRGPRWALSVTMSYLEI	179
Db	121	AKVVLAYGPTGAGKT-THAGQPRATGWDPAGSHGPPAAHKGGCGRGPMGLSVTMSYLEI	179
Qy	180	YQEKVLDLLDPASGDLVIREDCRGNILIPGLSOKPISFADPERHFLPASRNRVTGATRL	239
Db	180	YQEKVLDLLDPASGDLVIREDCRGNILIPGLSOKPISFADPERHFLPASRNRVTGATRL	239
Qy	240	NQRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGSEDNRRRTGNKGLRKESGAIN	299
Db	240	NQRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGSEDNRRRTGNKGLRKESGAIN	299
Qy	300	SLFVLGKVVDAINQGLPRVPYRDSKLTLLQDSLGGSASHSILIANIAPERFYLDVTVAL	359
Db	300	SLFALGKVVDAINQGLPRVPYRDSKLTLLQDSLGGSASHSILIANIAPERFYLDVTVAL	359
Qy	360	NFAARSKEVIN 370	
Db	360	NFAARSKEVIN 370	

C:Genetics:
A:Gene: FlyBase:K1P68D; KLP5
A:Cross-references: FlyBase:FBgn0004381
C:Function:
A:Description: may be part of a motor protein that provides anterograde fast axonal transport
C:Superfamily: kinesin-related protein KIF3; kinesin motor domain homology
C:Keywords: ATP; coiled coil; microtubule binding; nucleotide binding; P-loop
F:1-349/Domain: head globular #status predicted <HGL>
F:20-350/Domain: kinesin motor domain homology <KMOT>
F:106-113/Region: nucleotide-binding motif A (P-loop)
F:350-580/Domain: helical rod #status predicted <ROD>
F:581-784/Domain: tail globular #status predicted <TGL>
F:112/Binding site: ATP (lys) #status predicted

Query Match 29.6%; Score 555.5; DB 1; Length 784;
Best Local Similarity 39.5%; Pred. No. 2.4e-36;
Matches 140; Conservative 56; Mismatches 127; Indels 31; Gaps 10;

QY 36 TRRP-----PPARVRVAVLRPFVDVTAGASPPCV-----RGMDSCSLEIANW--RN 81
DB 5 SRPRTGSSQTPNECVQVVRCPMSNRERSRSPVVNVYVNRGV---VELQNVVDGN 60

QY 82 HOETLKYQDFAFYGRSRTQDIYAGSVOPILRHLEQNASVLAYGPTGACKTHTM--L 138
DB 61 KEQRKVFTYDAAYDASATQTTLYHEVFPFLVSSVLEGFNGCFYAGTGTGKTFTMEGVR 120

QY 139 GSPEQPGVTPRALMDL-LQLTREEGAEGRPALSVTMSYLEIYQEKVLDDLDPASGDIVI 197
DB 121 GNDELNGIIPRTFEQIWLHINTEN-----PQFLVDVSYLEIYMEELRDLKPKNSKHLEV 175

QY 198 REDCRN-ILIPGLSQKPISFPADFERHFLPASNRVTGATRLNORSRSHAVLLVKVDQ 256
DB 176 RE-RGSGVYVNLHAINCKSVEDMIKVMQGVNKRTVGTFTNMNEHSRSHAFIMKIE 233

QY 257 REELAPPROEGKVLIDLAGEDNRRRTGKGLRLEKSGAINTSLFVLGKVDALNQGLP 316
DB 234 CD-TEYNTTKVGNLIDLAGEDROSKTGASERLKEASKNLALSSLGNIWISALAESPP 292

QY 317 RVPYRDSKULTRLDPSLGSASHLSILIANIAPERRPYLDVTVSALNFAARSKVIN 370
DB 293 HVPYRDSKULTRLDPSLGSNSKTIIMANIGPSNYNYNETLTTLRYGSRKSIQN 346

RESULT 4
A86319
F25116.11 protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence revision 02-Mar-2001 #text change 09-Jul-2004
C:Accession: A86319
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Mailli, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A86319
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-703 <SFO>
A:Cross-references: UNIPROT.Q9PZ77; GB:AE005172; NID:g9795601; PIDN:AAF98419.1; GSPDB:GN C:Genetics:
A:Map position: 1

Query Match 29.3%; Score 550; DB 2; Length 703;
Best Local Similarity 38.6%; Pred. No. 5.6e-36;
Matches 130; Conservative 62; Mismatches 113; Indels 32; Gaps 6;
QY 42 ARVAVAVLRPFVDVTAGASPPCVCRGMDSCSLEIANWRNHQETLK-----YQDFAF 93

	Query Match	27.9%;	Score 524;	DB 2;	Length 1056;
	Best Local Similarity	36.4%;	Pred. NO. 1.2e-33;		
	Matches 131;	Conservative 67;	Mismatches 118;	Indels 44;	Gaps 9;
QY	44	VRVAVRLRFVDGTAGASDPPCCVRGMDSCSLELTANWRN	----	HQETILKYOFDAFYGERST	99
DB	13	VQVLLRCRPSDDEL-RSNAPQVLTCDQREAVVSQNTAGKHIDRV-FTFDKVGFSQAQ	70		
QY	100	QDDIYAGSVQPIRLRHLEGGONASVLAYGPTGAGKTHTMLGS	-----	PEQPGVI	147
DB	71	QKXDLYDAQVPIVNEVLEGFNCTIIPAYGQTGTGKTYTMEGECRRKSAPCGGLPAEAGVI	130		
QY	148	PRALMDLLOLQITREGAEGRPWALSVTMSYLEIYQEKVLDLLDP-	-----	ASGDV	196
DB	131	PRAVKQIF-----DTLEGQOAAYSVKVTFELNYEITDLLAPEDLSRVAABEKOKKPLP	185		

Qy	197	IREDQRGNILPGLSOKPSS-----FADFERHFLPASRENTVGATRLNORSSRSHAVILLV	255
Db	186	LMEDQGGVVLVRGLSEETVIGANEIYTLRL-----GSKRTAETFLNKOSSRSHLSFLSI	241
Qy	253	KYDQERLAPPRQ--REGKLYLIDLAGSEDRNRGTNGKGLRKSGGAINSTLFLVGLKVVDA	310
Db	242	THIKTEATPEGBELIKCGKMLNVLDSGSENTSRSGARDRGAREAGEINKSLLLTLGRVISA	301
Qy	311	LNQGLPVRVPRDSKTLRLQDLSLGGSAHSILIANIAPERRFYDLTVTSALNFAARSKEVIN	370
Db	302	LVHLGHVPRDSKTLRLRSLSGRTKTCIIATVSPAVHCLERTSLTLDYAHRAKNINR	361

RESULT 13

T02017
kinesin-related protein TKRP125 - common tobacco
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 09-Jul-2004
C:Accession: T02017
R:Asada, T.; Kuriyama, R.; Shibaoka, H.
J. Cell Sci. 110, 179-189, 1997
A:Title: TKRP125, a kinesin-related protein involved in the centrosome-independent organ
A:Reference number: Z14490; MUID:97136959; PMID:9044048
A:Accession: T02017
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1006 <ASA>
A:Cross-references: UNIPROT:Q23826; EMBL:D83711; NID:g2582970; PIDN:BAA23159.1; PID:g258
C:Genetics:
A:Note: TKRP125
C:Superfamily: kinesin-related protein E95; kinesin motor domain homology
E:10-361/Domain: kinesin motor domain homology <KMOT>

Query Match	27.8%	Score 521.5;	DB 2;	Length 1006;
Best Local Similarity	35.6%	Pred. No. 1.8e-33;		
Matches 128. Conservative	69.	Mismatches 118.	Indels 45.	Gaps 9;

QY	44	VRVAVRLRPFDVGTAGASDP	PCVRGMD-----SCSLEITANWRNHQETLKYOFDAFYGERS	98
DB	10	VOVLLRCRPFENDELURNAPQVVT	CNDYQREVAVSONIAG--KHIDRI-FTFDKVKFGESA	66
QY	99	TQODIYAGSVOPILRHLLLEGONAS	VLAYGPTCAGKHTMTLGS-----PEOPGV	146
DB	67	QORDLYDQAIVPINVEVLEGFNCTI	FAYGQTGTGKTYTMEGECKRSKSGPNGLPQEA	126
QY	147	IPRALMDLLQLTREGAEGRPWALS	VTMSYLEIYQEKVLDLLDP-----ASGLV	196
DB	127	IPRAVKQVFDLTLESQNAE-----	YSVKVTFLELYNEEITDLLAPEDLKVALEDQKKQLP	181
QY	197	TREDRCRGNILPGLSQKPISS----	FADFERHFLPASRNRTVGATRLNQRSRSRSHAVLLV	252
DB	182	LMEDGKGVLVRGEEIVTSANEFTLL	ER-----GSAKRRTAETLLNKQSSRSRSHLSFI	237
QY	253	KVDQERLAPFRQ--REGKLYLIDLAG	SEDNRTGNKGLUKESCAINTSLFVLGKVDA	310
DB	238	TTHIKEATPEGBELKCGKLNIVDLAG	SNSRSRGAREGAREGEINKSLTULGRVINA	297
QY	311	LNQOGLPVVPYRDSKLTLLIQSLIG	GSASHILIANIATAPERRFYDLTVSALNFAARSKEVIN	370
DB	298	IVENHCHTVPYRDSKLTLLIPDSI	GGRTKCTIATVSPAVHCEETLSLIDVAHRAKNIK	357

RESULT 14

probable kinesin-like spindle protein [imported] - *Arabidopsis thaliana*
 B94687
 C-Species: *Arabidopsis thaliana* (mouse-ear cress)
 C-Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
 C-Accession: B84687
 R-Rubin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Tallon, L.
 eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.B.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 Abstract: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:38:46 ; Search time 5.60933 seconds
(without alignments)
1056.470 Million cell updates/sec

Title: US-10-797-893-2
Perfect score: 1878
Sequence: 1 MAAGSTQORREMAASAA.....FYLDVTSAFNFAARSKEVIN 370

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 67802 seqs, 16016457 residues

Total number of hits satisfying chosen parameters: 67802

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New:
1: /cgn2_6/ptodata/2/paa/pct NEW COMB.pcp.*
2: /cgn2_6/ptodata/2/paa/US06 NEW COMB.pcp.*
3: /cgn2_6/ptodata/2/paa/US07 NEW COMB.pcp.*
4: /cgn2_6/ptodata/2/paa/US08 NEW COMB.pcp.*
5: /cgn2_6/ptodata/2/paa/US09 NEW COMB.pcp.*
6: /cgn2_6/ptodata/2/paa/US10 NEW COMB.pcp.*
7: /cgn2_6/ptodata/2/paa/US11 NEW COMB.pcp.*
8: /cgn2_6/ptodata/2/paa/US60 NEW COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	552.5	29.4	898	6	US-10-734-049A-259
2	526	28.0	1380	6	US-10-952-698-164
3	423	22.5	361	6	US-10-948-973-2
4	423	22.5	369	6	US-10-948-973-6
5	423	22.5	370	6	US-10-948-973-4
6	394	21.0	494	6	US-10-951-582-4
7	394	21.0	658	6	US-10-951-582-6
8	390	20.8	433	6	US-10-951-582-2
9	371	19.8	960	6	US-10-951-582-8
10	206.5	11.0	137	6	US-10-220-366A-19691
11	183	9.7	134	6	US-10-220-366A-14894
12	180	9.6	134	6	US-10-220-366A-26476
13	174	9.3	139	6	US-10-220-366A-15821
14	113	6.0	911	6	US-10-732-923-6841
15	113	6.0	911	6	US-10-732-923-6842
16	112	6.0	460	6	US-10-732-923-7092
17	103.5	5.5	913	6	US-10-732-923-6859
18	103.5	5.5	913	6	US-10-732-923-6860
19	99.5	5.3	911	6	US-10-732-923-6846
20	99.5	5.3	911	6	US-10-732-923-6847
21	96.5	5.1	912	6	US-10-732-923-6851
22	96.5	5.1	914	6	US-10-732-923-6854
23	96	5.1	465	6	US-10-732-923-6848
24	95.5	5.1	848	6	US-10-732-923-6904
25	95.5	5.1	877	6	US-10-732-923-6751

26	95	5.1	912	6	US-10-732-923-424	Sequence 424, App
27	95	5.1	912	6	US-10-732-923-7145	Sequence 7145, App
28	95	5.1	912	6	US-10-732-923-7146	Sequence 7146, App
29	95	5.1	912	6	US-10-732-923-7147	Sequence 7147, App
30	95	5.1	912	6	US-10-732-923-7148	Sequence 7148, App
31	95	5.1	912	6	US-10-732-923-7149	Sequence 7149, App
32	94.5	5.0	912	6	US-10-732-923-6852	Sequence 6852, App
33	94	5.0	852	6	US-10-732-923-6770	Sequence 6770, App
34	93.5	5.0	865	6	US-10-732-923-7188	Sequence 7188, App
35	93	5.0	875	6	US-10-732-923-7198	Sequence 7198, App
36	92.5	4.9	846	6	US-10-732-923-7199	Sequence 7199, App
37	91	4.8	909	6	US-10-732-923-6853	Sequence 6853, App
38	90.5	4.8	918	6	US-10-732-923-6855	Sequence 6855, App
39	90.5	4.8	918	6	US-10-732-923-6856	Sequence 6856, App
40	90	4.8	361	6	US-10-732-923-10271	Sequence 10271, A
41	89.5	4.8	900	6	US-10-732-923-7200	Sequence 7200, App
42	88.5	4.7	839	6	US-10-732-923-6936	Sequence 6936, App
43	88.5	4.7	878	6	US-10-732-923-6973	Sequence 6973, App
44	88.5	4.7	918	6	US-10-732-923-6857	Sequence 6857, App
45	88.5	4.7	918	6	US-10-732-923-6858	Sequence 6858, App

ALIGNMENTS

RESULT 1
US-10-734-049A-259
; Sequence 259, Application US/10734049A
; GENERAL INFORMATION:
; APPLICANT: ITOH, Kyogo
; TITLE OF INVENTION: TUMOR ANTIGEN
; FILE REFERENCE: Q-78382
; CURRENT APPLICATION NUMBER: US/10/734,049A
; PRIOR FILING DATE: 2003-12-12
; PRIOR APPLICATION NUMBER: PCT/JP02/05799
; PRIOR FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: JP 2001/177058
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: JP 2001/250728
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 408
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 259
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-734-049A-259

Query Match	29.4%	Score 552.5;	DB 6;	Length 898;
Best Local Similarity	39.7%	Pred. No. 1.1e-43;		
Matches 141;	Conservative 59;	Mismatches 114;	Indels 41;	Gaps 11;
QY	44	VRVAVRLRP-----FVDTAGADP-----PCVRGMDSCSLEIANWRHQ	83	
DB	12	MKVVRVPENTKEKAGFKHVHVDGHLVDFPKQEVSFHFGKKTNNQVIKQN---	69	
QY	84	ETLYQDFAFYGERSTQDIYAGSVQPIRLHLLSGQNASVLAYGPTGAGKTHMLGSPEQ	143	
DB	70	KDLKFVDAVDFDETSTQSEVFHTTKPLRSFLNGYNTVLAYGATGAGKTHMLGSADE	129	
QY	144	PGVTPRALMDLLQLTREGAEGRPWALSVTWMSYLEIYQEKVLDLLDPASGDIVIREDCRG	203	
DB	130	PGVMTLMLHYKCMDEIKEEK---ICSTAVSYLEVYNEQIRDLL--VNSGPLAVREDTQK	185	
QY	204	NILIPGLS-QKPISSPADFERHFLP-ASRNRVTGATRLNQRSSRSHAVLLVKVDQERLA	261	
DB	186	GVVHGLTLHQPKSS--BEILHLLDNGKNKNTQHTDMNATSSRSHAVFQIYLRQDKTA	243	
QY	262	PFRQ--REGKYLILDLAGSEDNRTGNKGLKSSGANTSLFLVGLKVVDAI-----NQ	313	
DB	244	SINQNVRIAKMSLIDLAGSERASTSGAKGTRFVEGTNINRSLLALGNVINALADSKRNQ	303	

```
Qy 314 GLPRVPYRDSKLTLLQDSLSGSAHSILIANIAPERFYLDITVSALNFAARKEV 368
Db 304 ---HIPYRNSKLTRELLKXSLGNGCQTIIMIAAVSPSVFYDDTYNTLKYANRAKDI 355

RESULT 2
US-10-952-698-164
; Sequence 164, Application US/10952698
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; TITLE OF INVENTION: METHODS OF DIAGNOSIS OF BLADDER CANCER. COMPOSITIONS AND METHODS
; TITLE OF INVENTION: OF SCREENING FOR MODULATORS OF BLADDER CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/952,698
; CURRENT FILING DATE: 2004-09-28
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 164
; LENGTH: 1380
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-952-698-164

Query Match 28.0%; Score 526; DB 6; Length 1380;
Best Local Similarity 37.2%; Pred. No. 6.3e-41;
Matches 128; Conservative 66; Mismatches 128; Indels 22; Gaps 7;

Qy 44 VRVAVRLRPFDV--GTAGASDPPCVRGWMDSCSLEIANRNHQTLYKQFDAYGERSTQ 101
Db 27 IKVFRIRPPAERSGSDAQEQNLCLSVLSSTSLRL---HSNPEKPTTFDHVADVDITQE 83

Qy 102 DIYAGSVQPIRLHLEGONASVLAYGPTGAGKTHMLGSPQEP-----GVIPRA---L 151
Db 84 SVFATVAKSIVSFCMSGVNGTIFAYGQTGSGKTFMTMGPSDFNSHNLGRGVIPIRSPEYL 143

Qy 152 MDLLQLFREEGACGRPWALSVTMSYLEIQEKVLDLDPASGDVIVREDCKGNILIPGLS 211
Db 144 FSLIDREKEKAGAGKSPFCK--CSFIEIYNEQIYDLDSAGLYLREHIKGVFVVGAV 201

Qy 212 QKPTISSFADPFERHFLPASRNRTVGATRLNQRSSRSHAVLVKVDORERLAPP--RQREGKL 270
Db 202 EQVTSAAEAQVLSGWRNRRVASTSNRRESSHVAFTTIESMEKSNEIVNIRTSLL 261

Qy 271 YLIDLAGSDNRRTGKGLRUKESGAINTSFLVGLKVVDAL----NOGLPRVPYRDSKLT 326
Db 262 NLVDLAGSERQKDTAEGMRLKEAGNINRSLSCLGQVITALVDVGNKGQRHVQYRDSKLT 321

Qy 327 RLQDSLSGSAHSILIANIAPERFYLDITVSALNFAARKEVIN 370
Db 322 FLRLDSLGNKARTAIANVHFGRCFGTLLTLNPAQRAKLKN 365

RESULT 3
US-10-948-973-2
; Sequence 2, Application US/10948973
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1043
; CURRENT APPLICATION NUMBER: US/10/948,973
; CURRENT FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: US/10/126,205
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US/09/723,595
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/295,612
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 361
```

```
; TYPE: PRT
; ORGANISM: Human
US-10-948-973-2

Query Match 22.5%; Score 423; DB 6; Length 361;
Best Local Similarity 32.8%; Pred. No. 4e-32;
Matches 121; Conservative 59; Mismatches 131; Indels 58; Gaps 11;

Qy 44 VRVAVRLRP-----FVDGTAGASDPPCVRGWMDSCSLEIANRNHQTLYK-- 87
Db 8 IRVFCRVRLPGEPTPPGGLLFPSPGPGSDPP-----TRLSRSDERRGTLSCA 60

Qy 88 -----YQDAFYGERSTQODIYAGSVQPIRLHLEGONASVLAYGPTGAGKTHML- 138
Db 61 PAPPTRHDFSFRVPVPGSGQDEVEF-EIAMLVQSALDGYPVCI FAYGQTGSGKFTTMEG 119

Qy 139 ---GSPEQPGVIPRALMDLLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDPAS---- 192
Db 120 GPGGDPQLEGLIPRALRHLSVAQELSGQG--WTYSFVASVYVEIYNETVRDLLATGTGK 177

Qy 193 ---GDLVIREDCRGN--ILIPGLSQKPISSFADPFERHFLPASRNRTVGATRLNQRSSRSH 248
Db 178 QGGECEIRRAGPGSEELVTNARYVPVSCKEVDALLHLARONRAVARTAQNERSRSHS 237

Qy 249 VLLVKVDORERLAPPFQREGKLYLIDLAGSDNRRTGNKGL-----RLKESGAINTS 300
Db 238 VFQLQI-SGEHSRGLQCGAPLSLVDLAGSE---RLDPGLALGPGERERLRETQAINSS 292

Qy 301 LFVLGKVDALNOGLPRVPYRDSKLTTRLLQDSLSGSAHSILIANIAPERFYLDITVSALN 360
Db 293 LSTGLGLVINALSNKESHVYRNSKLTLYLQNSLGGSAKMLFVNISPLEENVSESLNSLR 352

Qy 361 FAARKEVIN 369
Db 353 FASKVNQCV 361

RESULT 4
US-10-948-973-6
; Sequence 6, Application US/10948973
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1043
; CURRENT APPLICATION NUMBER: US/10/948,973
; CURRENT FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: US/10/126,205
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US/09/723,595
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/295,612
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Human
US-10-948-973-6

Query Match 22.5%; Score 423; DB 6; Length 369;
Best Local Similarity 32.8%; Pred. No. 4.1e-32;
Matches 121; Conservative 59; Mismatches 131; Indels 58; Gaps 11;

Qy 44 VRVAVRLRP-----FVDGTAGASDPPCVRGWMDSCSLEIANRNHQTLYK-- 87
Db 7 IRVFCRVRLPGEPTPPGGLLFPSPGPGSDPP-----TRLSRSDERRGTLSCA 59

Qy 88 -----YQDAFYGERSTQODIYAGSVQPIRLHLEGONASVLAYGPTGAGKTHML- 138
Db 60 PAPPTRHDFSFRVPVPGSGQDEVEF-EIAMLVQSALDGYPVCI FAYGQTGSGKFTTMEG 118
```

```
QY 139 ---GSPEQGVIPRALMDLLQLTREBAGRPPWALSVTWMSYLEIYOEKVLDDLPAS--- 192
Db 119 GPGGDPQLEGLIPRALRHFLFVAQELSGQG--WTYSFVASVVEIYNETVRDLLATGTRKG 176
QY 193 --GDLVIREDCRGN--ILTPGLSQKPISSFADFERHFLPASRRTVTGATRLNQRSSRSHA 248
Db 177 QGGECEIRRAGSGSELVTNARYVPVSCKEVDALLHLARQNAVARTAQNRSRSHS 236
QY 249 VLLVKVDQERLAPRQREGKLYLIDLAGESENRRTGNKGL-----RLKESGAINTS 300
Db 237 VFQLOI-SGEHSSRGLQCGAPLSLVDLAGE-----RLDPGLALGPERERLRETQAINSS 291
QY 301 LFLVGLKVVDALNOGLPRVPYRDSKTLRLQDSLGSAHSILIANIAPERRFYLDTVSALN 360
Db 292 LSTLGLVIMALSNKESHVPYRNSKLTLYLLQNSLGGSAKMLMFVNISPLEENVSESLNSLR 351
QY 361 FAARSKEVI 369
Db 352 FASKVNOCV 360

RESULT 5
US-10-948-973-4
; Sequence 4, Application US/10948973
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1043
; CURRENT APPLICATION NUMBER: US/10/948,973
; CURRENT FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: US/10/126,205
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US/09/723,595
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/295,612
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Human
US-10-948-973-4

Query Match 22.5%; Score 423; DB 6; Length 370;
Best Local Similarity 32.8%; Pred. No. 4.1e-32;
Matches 121; Conservative 59; Mismatches 131; Indels 58; Gaps 11;

QY 44 VRVAVRLRP-----FVDGTAGASDPPCVRGMDSCSLEIANWRNHQETLK-- 87
Db 8 IRVFCVRVPLGPEPTPPGLLLFPSPGPGSDPP-----TRLSSRSDERRGTLGSA 60
QY 88 -----YQFDAFYGERSTQDDIYAGSVQPIRLHLEGONASVLAYGPTGAGKTHML- 138
Db 61 PAPTRHDFSDVRFPFGSGQDEVE-BIAMLVQSGALDGYPCIPAYGQTGSGKTFMEG 119
QY 139 ---GSPEQGVIPRALMDLLQLTREBAGRPPWALSVTWMSYLEIYOEKVLDDLPAS--- 192
Db 120 GPGGDPQLEGLIPRALRHFLFVAQELSGQG--WTYSFVASVVEIYNETVRDLLATGTRKG 177
QY 193 --GDLVIREDCRGN--ILTPGLSQKPISSFADFERHFLPASRRTVTGATRLNQRSSRSHA 248
Db 178 QGGECEIRRAGSGSELVTNARYVPVSCKEVDALLHLARQNAVARTAQNRSRSHS 237
QY 249 VLLVKVDQERLAPRQREGKLYLIDLAGESENRRTGNKGL-----RLKESGAINTS 300
Db 238 VFQLOI-SGEHSSRGLQCGAPLSLVDLAGE-----RLDPGLALGPERERLRETQAINSS 292
QY 301 LFLVGLKVVDALNOGLPRVPYRDSKTLRLQDSLGSAHSILIANIAPERRFYLDTVSALN 360
Db 293 LSTLGLVIMALSNKESHVPYRNSKLTLYLLQNSLGGSAKMLMFVNISPLEENVSESLNSLR 352
```

```
QY 361 FAARSKEVI 369
Db 353 FASKVNOCV 361
```

RESULT 6

```
US-10-951-582-4
; Sequence 4, Application US/10951582
; GENERAL INFORMATION:
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1041
; CURRENT APPLICATION NUMBER: US/10/951,582
; CURRENT FILING DATE: 2004-09-27
; PRIOR APPLICATION NUMBER: US/10/143,563
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: US/09/723,630
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/595,424
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/314,464
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Human
US-10-951-582-4
```

```
Query Match 21.0%; Score 394; DB 6; Length 494;
Best Local Similarity 27.3%; Pred. No. 3.3e-29;
Matches 117; Conservative 64; Mismatches 127; Indels 120; Gaps 13;
```

```
QY 44 VRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIAN-----WR-----NHQETLK 87
Db 24 VGVYCRVRL-----GFPD-----QECCEVINTVTQLHTPEGYRLNRNGDKRET-Q 70
QY 88 YQFDAFYGERSTQDDIYAGSVQPIRLHLEGONASVLAYGPTGAGKTHMLGSPQPGVI 147
Db 71 YSFKQVFGTHTQKELPDVANPLVNDLIHGKNGLLFTYGVGTGSKTHMTGSPGEGILL 130
QY 148 PRAL-----MD-----LLQLTREBAGRPPWALS----- 171
Db 131 PRCLDMIFNSIGSFQAKRYVFKSNDNRMSMDIQCEVDALLERQKREAMPNPKTSSKRQVD 190
QY 172 -----VTWMSYLEIYOEKVLDDLPASGDLV-----I 197
Db 191 PEPADMITVQFCAEVEDSDSVGVFVSYIETIYNNYIYDLLEEVFPDPIKPKPQSKLL 250
QY 198 REDCRGNILPGLSQKPISSFADFERHFLPASRRTVTGATRLNQRSSRSHAVLLVKVDQR 257
Db 251 REDKNHNNYVAGTEVEVKSTEEAFEFWRGQKGRRIANTHLNRESSRSHSVFNKLVQ- 309
QY 258 ERLAP-----PRORE-----GKLYLIDLAGESENRRTGNKGLRKESGAINTSFLVUG 305
Db 310 ---APLDADGDNVLQKEQKITISQLSLVDLAGSERTNTRAEGNRLREAGNINQSLMTLR 366
QY 306 KVDALNOGL-----PRVPYRDSKTLRLQDSLGSAHSILIANIAPERRFYLDTVSALN 360
Db 367 TCMVDVLENQMYTNKVPYRDSKTLHLFKNYFDGEGKVRMIVCVNPKAEDYEENLQVMR 426
QY 361 FAARSKEV 368
Db 427 FAEVTQEV 434
```

```
RESULT 7
US-10-951-582-6
; Sequence 6, Application US/10951582
```

```

; GENERAL INFORMATION:
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1041
; CURRENT APPLICATION NUMBER: US/10/951,582
; CURRENT FILING DATE: 2004-09-27
; PRIOR APPLICATION NUMBER: US/10/143,563
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: US/09/723,630
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US/09/595,424
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/314,464
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Human
; US-10-951-582-6

Query Match      21.0%; Score 394; DB 6; Length 658;
Best Local Similarity 27.3%; Pred. No. 5e-29;
Matches 117; Conservative 64; Mismatches 127; Indels 120; Gaps 13;

QY 44 VRVAVRLRPFDGTAGASDPPCVRGMDSCSLEIAN-----WR-----NHQETLK 87
DB 24 VGVYCRVRPL-----GFPD-----QECIEVINNTVQLHTPEGYRLNRNGDYKET-Q 70

QY 88 YQFDAYGERSTOODIYAGSVQPIRLHLEQONASVLAYGPTGAGKTHTMLGSPQPGVI 147
DB 71 YSPKQVGTHTTQKELFDVVANPLVNDLHGKNGLLFTYGVGTGSKTHMTGSPGEGGL 130

QY 148 PRAL-----MD-----LLQLTREEGAGRPWALS-----171
DB 131 PRCLDMIFNSIGSFQAKRYVFKSNDNRNMDIQCEVDALLERQKREAMPNKTSSKRQVD 190

QY 172 -----VTMSYLEIQEKVLDLLDPASGDLV-----I 197
DB 191 PEFADMITVQEFKAEVEDESDSVGVFVSIEIYNNYIYDLLEVPDPKPKPPQSKLL 250

QY 198 REDCRGNILPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQR 257
DB 251 REDKNHNMVAGTEVEVKSTEEAFVFWRGQKKRIANTHLNRESRSSHVSFNKLVQ- 309

QY 258 ERLAP-----FQRE-----GKLYLIDLAGSEDNRRTGNKGLRLKESGAINTSFLVLG 305
DB 310 ---APLDADGDNVLQEKQITISQLSLVDLAGSERTNRTAEGNRLREAGNIQSLMTLR 366

QY 306 KVDALNQGGL-----PRVPRYRDSKLTRLLODSLGSSAHSILIANIAPERFYLDITVSALN 360
DB 367 TCMVDVLRNQMVGTKMVPYRDSKLTFLFKNYPDGEKGKVRMIVCVNPKAEDYEENLQVMR 426

QY 361 FAARSKE 367
DB 427 FAEVTOE 433

RESULT 9
US-10-951-582-8
; Sequence 8, Application US/10951582
; GENERAL INFORMATION:
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1041
; CURRENT APPLICATION NUMBER: US/10/951,582
; CURRENT FILING DATE: 2004-09-27
; PRIOR APPLICATION NUMBER: US/10/143,563
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: US/09/723,630
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US/09/595,424
; PRIOR FILING DATE: 2000-06-15
; CURRENT APPLICATION NUMBER: US 09/314,464

```

[illegible]

```

RESULT 12
US-10-220-366A-26476
; Sequence 26476, Application US/10220366A
; GENERAL INFORMATION:
; APPLICANT: HYSEQ, INC
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-042
; CURRENT APPLICATION NUMBER: US/10/220,366A
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/577,409
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/515,126
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 27802
; SOFTWARE: Custom
; SEQ ID NO 26476
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-220-366A-26476

Query Match          9.6%; Score 180; DB 6; Length 134;
Best Local Similarity 33.3%; Pred. No. 5.6e-10;
Matches 42; Conservative 21; Mismatches 53; Indels 10; Gaps 2

QY 41 PAR-----VRVAVRLRPFFVDGTAGASDPPCVKRGDMSCSLEIANRNHNHETLKY-----QF 90
Db 1 PAREMEKFRVCVRKEPLGMRREVGRGEINIITVEDKETLLVHKEKEAVDLTQYILQHVFF 60

QY 91 DAFYGERSTQODIVAGSVQPIRLHLLSEGNQNASVLAYGPTGAGKTHTMLGSPGFGVPIRA 150
Db 61 DEVFGEACTNODVYMKTHPLIOHIFNGGNATCFAYGQTGAGKTYMIGTHENPGLYALA 130

```

```
QY 151 LMDLLQ 156
Db 121 AKUIFR 126

RESULT 13
US-10-220-366A-15821
; Sequence 15821, Application US/10220366A
; GENERAL INFORMATION:
; APPLICANT: HYSEQ, INC
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-042
; CURRENT APPLICATION NUMBER: US/10/220,366A
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/577,409
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/515,126
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 27802
; SOFTWARE: Custom
; SEQ ID NO 15821
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; LOCATION: (1)-(139)
; OTHER INFORMATION: Xaa = any amino acid or nothing
US-10-220-366A-15821

Query Match 9.3%; Score 174; DB 6; Length 139;
Best Local Similarity 43.2%; Pred. No. 2.1e-09;
Matches 41; Conservative 17; Mismatches 25; Indels 12; Gaps 2;

QY 288 GIRLKESAINTSFLVLGKVVDALE---NOGLPR-----VYRDSKLTLLQLDSLGG 335
Db 6 GVRLKEGKNITESTFVTGNGVTSALADFSQDASRYSKKKQVLVYRDSVMTWLLKDSLGG 65

QY 336 SAHSILIANIAPERFRFYLDITVSALNFAARKEVIN 370
Db 66 NSKALMWTATISPADWNYGETLSTLYANRAKNIIN 100

RESULT 14
US-10-732-923-6841
; Sequence 6841, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 6841
; LENGTH: 911
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-732-923-6841

Query Match 6.0%; Score 113; DB 6; Length 911;
Best Local Similarity 21.4%; Pred. No. 0.018;
Matches 83; Conservative 69; Mismatches 128; Indels 108; Gaps 20;

QY 2 AAGGSTQORRREMAAASAAAGRCRLSKIGATRRPP---PA-----RVRVAVR 49
Db 52 SAGG-----ENAAQSAERVINQA---LKKLPSPPPDDIPASSSLIKVIRRAQAQK 101

QY 50 LRP-----FVD-----GTAGA---SDPPCVRGMDSCSLEIANWRNHQ 83
Db 102 SRGDTHLAVDQLIMGLLEDSQIRDLLNEVGAVATARKSEVEKLRGKGVESASGDTNF 161

QY 84 ETLKYQFDAYGERSQODIYAGSVQPII-----RHLLEGQNASVLAYGPTGAGK 133
Db 162 QALK-----TYGRDLVEQ---AGKLDPVIGRDEIRRVVRLSRTYKNPVLIGEPGVGK 213

QY 134 THTMLGSPQEQ--PGVIPRALMDLLQLTREGAERGPWALSVTMSYLEIYQEK---VLDLL 188
Db 214 TAVVEGLAQRIKVGDVFNLSLTDVRLISLDMG-----ALVAGAKYRGFEERLKSVLKEV 267

QY 189 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNRVTGATRLNQRSSSHA 248
Db 268 EDAGKVLIFIDEIHLVLGAGKTEGSMDA-ANLFPKMLARGQLRCIGATTLEE--YRKV 324

QY 249 VLLVKVQDERLAPFRQREGKLYLD-----LAGSEDRRTTGNKGLRKESGAIN- 298
Db 325 -----EKDAAFERRFOQVVAEFSVPDTISILGLKE-KYEGHHGVRIQDRALINA 374

QY 299 ---TSLFVLG-----KVVDALNQGLPRV 318
Db 375 AQLSARYITGRHLDPDKAIDLVDACANV 402

RESULT 15
US-10-732-923-6842
; Sequence 6842, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 6842
; LENGTH: 911
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-732-923-6842

Query Match 6.0%; Score 113; DB 6; Length 911;
Best Local Similarity 21.4%; Pred. No. 0.018;
Matches 83; Conservative 69; Mismatches 128; Indels 108; Gaps 20;

QY 2 AAGGSTQORRREMAAASAAAGRCRLSKIGATRRPP---PA-----RVRVAVR 49
Db 52 SAGG-----ENAAQSAERVINQA---LKKLPSPPPDDIPASSSLIKVIRRAQAQK 101

QY 50 LRP-----FVD-----GTAGA---SDPPCVRGMDSCSLEIANWRNHQ 83
Db 102 SRGDTHLAVDQLIMGLLEDSQIRDLLNEVGAVATARKSEVEKLRGKGVESASGDTNF 161

QY 84 ETLKYQFDAYGERSQODIYAGSVQPII-----RHLLEGQNASVLAYGPTGAGK 133
Db 162 QALK-----TYGRDLVEQ---AGKLDPVIGRDEIRRVVRLSRTYKNPVLIGEPGVGK 213

QY 134 THTMLGSPQEQ--PGVIPRALMDLLQLTREGAERGPWALSVTMSYLEIYQEK---VLDLL 188
Db 214 TAVVEGLAQRIKVGDVFNLSLTDVRLISLDMG-----ALVAGAKYRGFEERLKSVLKEV 267

QY 189 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNRVTGATRLNQRSSSHA 248
Db 268 EDAGKVLIFIDEIHLVLGAGKTEGSMDA-ANLFPKMLARGQLRCIGATTLEE--YRKV 324

QY 249 VLLVKVQDERLAPFRQREGKLYLD-----LAGSEDRRTTGNKGLRKESGAIN- 298
Db 325 -----EKDAAFERRFOQVVAEFSVPDTISILGLKE-KYEGHHGVRIQDRALINA 374

QY 299 ---TSLFVLG-----KVVDALNQGLPRV 318
Db 375 AQLSARYITGRHLDPDKAIDLVDACANV 402
```

Search completed: November 5, 2004, 19:04:48
Job time : 6.60933 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:44:52 ; Search time 62.7813 Seconds
(without alignments)
2081.726 Million cell updates/sec

Title: US-10-797-893-2

Perfect score: 1878

Sequence: 1 MAAGSGTQORREMAAASAA.....FYLDTSALNFAARSKEVIN 370

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*

2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*

3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*

4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*

5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*

6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*

7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*

8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*

9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*

10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*

11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*

12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*

13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*

14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*

15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*

16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*

17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*

18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*

19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1878	100.0	370	16 US-10-797-893-2	Sequence 2, Appli
2	1878	100.0	490	15 US-10-334-143-80	Sequence 80, Appli
3	1873	99.7	512	16 US-10-797-893-4	Sequence 4, Appli
4	1763	93.9	346	16 US-10-797-893-6	Sequence 6, Appli
5	1763	93.9	460	9 US-09-925-300-1228	Sequence 1228, Ap
6	1763	93.9	487	16 US-10-797-893-8	Sequence 8, Appli
7	601	32.0	589	17 US-10-425-115-320209	Sequence 320209,
8	597.5	31.8	377	16 US-10-437-963-159031	Sequence 159031,
9	559	29.8	548	15 US-10-108-260A-2692	Sequence 2692, Ap
10	557.5	29.7	783	17 US-10-425-115-332919	Sequence 332919,
11	557.5	29.7	854	15 US-10-425-114-59708	Sequence 59708, A
12	544.5	29.0	776	16 US-10-437-963-103943	Sequence 103943,
13	528	28.1	383	9 US-09-883-096-5	Sequence 5, Appli

14	528	28.1	864	9 US-09-883-096-2	Sequence 2, Appli
15	526	28.0	375	15 US-10-332-089-4	Sequence 4, Appli
16	526	28.0	409	15 US-10-332-089-6	Sequence 6, Appli
17	526	28.0	774	14 US-10-369-493-2361	Sequence 2361, Ap
18	526	28.0	1388	14 US-10-146-473-82	Sequence 82, Appli
19	526	28.0	1388	15 US-10-173-999-32	Sequence 32, Appli
20	526	28.0	1388	15 US-10-332-089-2	Sequence 2, Appli
21	526	28.0	1388	15 US-10-188-832-164	Sequence 164, App
22	525.5	28.0	757	15 US-10-220-120-366	Sequence 366, App
23	525.5	28.0	757	16 US-10-363-829-316	Sequence 316, App
24	525	28.0	1030	15 US-10-443-114-62748	Sequence 62748, A
25	523.5	27.9	834	15 US-10-287-226-306	Sequence 306, App
26	521.5	27.8	1401	15 US-10-287-226-142	Sequence 142, App
27	520	27.7	1001	17 US-10-425-115-231231	Sequence 231231,
28	519.5	27.7	1826	16 US-10-408-765A-1477	Sequence 1477, Ap
29	518.5	27.6	1382	16 US-10-437-963-176714	Sequence 176714,
30	518	27.6	1115	15 US-10-287-226-302	Sequence 302, App
31	518	27.6	1232	14 US-10-116-712-670	Sequence 670, App
32	518	27.6	1232	16 US-10-408-765A-2153	Sequence 2153, Ap
33	518	27.6	1235	15 US-10-334-143-8	Sequence 8, Appli
34	517.5	27.6	1324	15 US-10-287-226-314	Sequence 314, App
35	516	27.5	992	17 US-10-739-930-7903	Sequence 7903, Ap
36	513.5	27.3	805	14 US-10-369-493-21903	Sequence 21903, A
37	513	27.3	1237	15 US-10-334-143-33	Sequence 33, Appli
38	510	27.2	672	16 US-10-408-765A-1664	Sequence 1664, Ap
39	509	27.1	1103	9 US-09-847-874A-1	Sequence 1, Appli
40	509	27.1	1103	14 US-10-458-162-1	Sequence 1, Appli
41	507	27.0	928	14 US-10-080-608A-23	Sequence 23, Appli
42	507	27.0	928	14 US-10-370-685-112	Sequence 112, App
43	507	27.0	1232	14 US-10-116-712-664	Sequence 664, App
44	507	27.0	1232	14 US-10-116-712-669	Sequence 669, App
45	504	26.8	338	9 US-09-883-096-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-10-797-893-2

; Sequence 2, Application US/10797893

; Publication No. US20040142397A1

; GENERAL INFORMATION:

; APPLICANT: Beraud, Christophe

; TITLE OF INVENTION: Novel motor proteins and methods for

; FILE OF INVENTION: their use

; FILE REFERENCE: 1044

; CURRENT APPLICATION NUMBER: US/10797,893

; PRIOR FILING DATE: 2004-03-09

; PRIOR APPLICATION NUMBER: US/09/724,224

; PRIOR FILING DATE: 2000-11-28

; PRIOR APPLICATION NUMBER: 09/597,292

; PRIOR FILING DATE: 2000-06-20

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 370

; TYPE: PRT

; ORGANISM: Human

US-10-797-893-2

Query Match	100.0%	Score	1878;	DB	16;	Length	370;
Best Local Similarity	100.0%	Pred. No.	4.3e-177;				
Matches	370;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MAAGSGTQORREMAAASAA	ISGAGRCRLSKIGATRRPPPARVAVRLRPVDTGTA	60			
Db	1	MAAGSGTQORREMAAASAA	ISGAGRCRLSKIGATRRPPPARVAVRLRPVDTGTA	60			
QY	61	SDPPCVRGMDSCSLEIANRNHSTLYKQFPAFYGERSTQODIYAGSVQPIRLHLEGN	120				
Db	61	SDPPCVRGMDSCSLEIANRNHSTLYKQFPAFYGERSTQODIYAGSVQPIRLHLEGN	120				
QY	121	ASVLAYGTGAGTKHTMLGSPQPGVPRALMDLLQLTREGAEGRPWALSVTMSYLETY	180				

Db 121 ASVLAYPTGAGKTHMLGSPQGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIY 180
QY 181 QEKVLDLDPASGDLVIREDCRGNILIPGLSQKPISSFADFERRHFLPASRNRVTGATRLN 240
Db 181 QEKVLDLDPASGDLVIREDCRGNILIPGLSQKPISSFADFERRHFLPASRNRVTGATRLN 240
QY 241 QRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGESEDRNRRTGNKGLRKESGAINTS 300
Db 241 QRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGESEDRNRRTGNKGLRKESGAINTS 300
QY 301 LFLVGKVVADALNOGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERFYLDTVSALN 360
Db 301 LFLVGKVVADALNOGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERFYLDTVSALN 360
QY 361 FAARSKEVIN 370
Db 361 FAARSKEVIN 370
RESULT 2
US-10-334-143-80
; Sequence 80, Application US/10334143
; Publication No. US20040009549A1
; GENERAL INFORMATION:
; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD
; FILE REFERENCE: 038602/1543
; CURRENT APPLICATION NUMBER: US/10/334,143
; PRIOR FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: 60/343,169
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 80
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-334-143-80
Query Match 100.0%; Score 1878; DB 15; Length 490;
Best Local Similarity 100.0%; Pred. No. 6.4e-177; Indels 0; Gaps 0;
Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAGGSTQORREMAAASAAAIISGAGRCRLSKIGATRRPPPARVRVAVRLRPPFVDGTAGA 60
Db 9 MAAGGSTQORREMAAASAAAIISGAGRCRLSKIGATRRPPPARVRVAVRLRPPFVDGTAGA 68
QY 61 SDPPCVRGMDSCSLEIANWRNHQETLKYQFDAPFYGERSTOODIYAGSVQPIRLHLLGQNA 120
Db 69 SDPPCVRGMDSCSLEIANWRNHQETLKYQFDAPFYGERSTOODIYAGSVQPIRLHLLGQNA 128
QY 121 ASVLAYPTGAGKTHMLGSPQGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIY 180
Db 129 ASVLAYPTGAGKTHMLGSPQGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIY 188
QY 181 QEKVLDLDPASGDLVIREDCRGNILIPGLSQKPISSFADFERRHFLPASRNRVTGATRLN 240
Db 189 QEKVLDLDPASGDLVIREDCRGNILIPGLSQKPISSFADFERRHFLPASRNRVTGATRLN 248
QY 241 QRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGESEDRNRRTGNKGLRKESGAINTS 300
Db 249 QRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGESEDRNRRTGNKGLRKESGAINTS 308
QY 301 LFLVGKVVADALNOGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERFYLDTVSALN 360
Db 309 LFLVGKVVADALNOGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERFYLDTVSALN 368
QY 361 FAARSKEVIN 370
Db 369 FAARSKEVIN 378

RESULT 3
US-10-797-893-4
; Sequence 4, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/797,893
; PRIOR FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Human
US-10-797-893-4
Query Match 99.7%; Score 1873; DB 16; Length 512;
Best Local Similarity 100.0%; Pred. No. 2.2e-176; Indels 0; Gaps 0;
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 AAGGSTQORREMAAASAAAIISGAGRCRLSKIGATRRPPPARVRVAVRLRPPFVDGTAGA 61
Db 3 AAGGSTQORREMAAASAAAIISGAGRCRLSKIGATRRPPPARVRVAVRLRPPFVDGTAGA 62
QY 62 DPCCVRGMDSCSLEIANWRNHQETLKYQFDAPFYGERSTOODIYAGSVQPIRLHLLGQNA 121
Db 63 DPCCVRGMDSCSLEIANWRNHQETLKYQFDAPFYGERSTOODIYAGSVQPIRLHLLGQNA 122
QY 122 SVLAYPTGAGKTHMLGSPQGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIY 181
Db 123 SVLAYPTGAGKTHMLGSPQGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIY 182
QY 182 EKVLDLDPASGDLVIREDCRGNILIPGLSQKPISSFADFERRHFLPASRNRVTGATRLN 241
Db 183 EKVLDLDPASGDLVIREDCRGNILIPGLSQKPISSFADFERRHFLPASRNRVTGATRLN 242
QY 242 RSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGESEDRNRRTGNKGLRKESGAINTS 301
Db 243 RSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGESEDRNRRTGNKGLRKESGAINTS 302
QY 302 FVLGKVVADALNOGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERFYLDTVSALN 361
Db 303 FVLGKVVADALNOGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERFYLDTVSALN 362
QY 362 AARSKEVIN 370
Db 363 AARSKEVIN 371
RESULT 4
US-10-797-893-6
; Sequence 6, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/797,893
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20

```

; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Human
US-10-797-893-6

Query Match          93.9%; Score 1763; DB 16; Length 346;
Best Local Similarity 100.0%; Pred. No. 9.7e-166;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 GRCLSKIGATRRPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 85
DB 2 GRCLSKIGATRRPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 61
;
QY 86 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQ 145
DB 62 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQ 121
;
QY 146 VIPRALMDLLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGNI 205
DB 122 VIPRALMDLLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGNI 181
;
QY 206 LIPGLSQPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 265
DB 182 LIPGLSQPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 241
;
QY 326 TRLLQDSLGSSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 370
DB 302 TRLLQDSLGSSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 346
;

RESULT 5
US-09-925-300-1228
; Sequence 1228, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1228
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (75)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (147)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (435)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1228

Query Match          93.9%; Score 1763; DB 9; Length 460;
Best Local Similarity 99.4%; Pred. No. 1.5e-165;
Matches 345; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 26 GRCLSKIGATRRPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 85
DB 2 GRCLSKIGATRRPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 61
;
QY 86 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQ 145
DB 62 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQ 121
;
QY 146 VIPRALMDLLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGNI 205
DB 122 VIPRALMDLLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGNI 181
;
QY 206 LIPGLSQPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 265
DB 182 LIPGLSQPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 241
;
QY 326 TRLLQDSLGSSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 370
DB 302 TRLLQDSLGSSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 346
;

RESULT 6
US-10-797-893-8
; Sequence 8, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/797,893
; PRIOR FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Human
US-10-797-893-8

Query Match          93.9%; Score 1763; DB 16; Length 487;
Best Local Similarity 100.0%; Pred. No. 1.6e-165;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 GRCLSKIGATRRPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 85
DB 2 GRCLSKIGATRRPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 61
;
QY 86 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQ 145
DB 62 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQ 121
;
QY 146 VIPRALMDLLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGNI 205
DB 122 VIPRALMDLLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGNI 181
;
QY 206 LIPGLSQPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 265
DB 182 LIPGLSQPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 241
;
QY 266 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVVDALNOGLPRVPYRDSKL 325
DB 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVVDALNOGLPRVPYRDSKL 301
;

```

```

; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 159031
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(377)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_58448C.1.pep
; US-10-437-963-159031

Query Match      31.8%; Score 597.5; DB 16; Length 377;
Best Local Similarity 39.7%; Pred. No. 3.7e-50;
Matches 138; Conservative 69; Mismatches 112; Indels 29; Gaps 8;

QY 30 LSKIGATRRPPPARVRVAVRLRPFDGTAGASDPPCVRGMDSCSLEIATNRNHOETLK- 87
Db 30 LSKIGATRRPPPARVRVAVRLRPFDGTAGASDPPCVRGMDSCSLEIATNRNHOETLK- 87
QY 1 MATAAATQSQP---VRVLRVRPHLPSEANGAACPVGLLGSHPGGEVTVQLKDYTSRN 57
Db 1 MATAAATQSQP---VRVLRVRPHLPSEANGAACPVGLLGSHPGGEVTVQLKDYTSRN 57
QY 88 --YQFDAPFGERSTQDDIYAGSVQPIRLHLLGONASVLAAYGPTGAGKTHMLGSPQPG 145
Db 88 --YQFDAPFGERSTQDDIYAGSVQPIRLHLLGONASVLAAYGPTGAGKTHMLGSPQPG 145
QY 58 ECKYLDAPFGQESRVECFDQESAVIFGIFEGTNATVAYATGSGKTYTMOGTEDLPG 117
Db 58 ECKYLDAPFGQESRVECFDQESAVIFGIFEGTNATVAYATGSGKTYTMOGTEDLPG 117
QY 146 VIPRALMDLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRNI 205
Db 146 VIPRALMDLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRNI 205
QY 118 LMKSTVIALCTGT-----WC-SVEISYYEVTMERCYDLLEPKAREIMVLDKGNL 167
Db 118 LMKSTVIALCTGT-----WC-SVEISYYEVTMERCYDLLEPKAREIMVLDKGNL 167
QY 206 LIPGUSQKPISSAFERHFLPASNRVTGATRLNQRSSRSHAVLLVKVDQERLAPFRQ 265
Db 206 LIPGUSQKPISSAFERHFLPASNRVTGATRLNQRSSRSHAVLLVKVDQERLAPFRQ 265
QY 168 QLKGLAWVPVRSLEEFHEIYSGVQRKVAHTGLNDVSSRSHAVLSIRTT-----DV 220
Db 168 QLKGLAWVPVRSLEEFHEIYSGVQRKVAHTGLNDVSSRSHAVLSIRTT-----DV 220
QY 266 REGKLYLI---DLAGSDNRRTGNKGLRLKESGAINTSFLVGLKVVDALNQLGPRVPYRD 322
Db 266 REGKLYLI---DLAGSDNRRTGNKGLRLKESGAINTSFLVGLKVVDALNQLGPRVPYRD 322
QY 221 VEGKMLNLTSLXDLAGNEDNRRTCNIGIRLOESAKINQSLFALSINVISALNKKEPRIPYRE 280
Db 221 VEGKMLNLTSLXDLAGNEDNRRTCNIGIRLOESAKINQSLFALSINVISALNKKEPRIPYRE 280
QY 323 SKITRLQDSLGSAHSILIANIAPERRFYLDVTSALNFAARKEVIN 370
Db 323 SKITRLQDSLGSAHSILIANIAPERRFYLDVTSALNFAARKEVIN 370
QY 281 SKLTRILQDSLGNSHAVMIACINPVE--YQEAHVTVSLAARSRHVTN 326
Db 281 SKLTRILQDSLGNSHAVMIACINPVE--YQEAHVTVSLAARSRHVTN 326

RESULT 9
US-10-108-260A-2692
; Sequence 2692, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cdna
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108, 260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2692
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-108-260A-2692

Query Match      29.8%; Score 559; DB 15; Length 548;
Best Local Similarity 38.9%; Pred. No. 4.2e-46;
Matches 132; Conservative 70; Mismatches 121; Indels 16; Gaps 7;

QY 43 RRVAVRLRPFF--VDGTAGAS-----DPPCVRGMDSCSLEIATNRNHOETLK-YQFADF 93
Db 43 RRVAVRLRPFF--VDGTAGAS-----DPPCVRGMDSCSLEIATNRNHOETLK-YQFADF 93
QY 11 QLMVALVRPISVAELBEGATLIAHKVDEQVMVMDPMDPDILRAHRSREKSYLFDVA 70
Db 11 QLMVALVRPISVAELBEGATLIAHKVDEQVMVMDPMDPDILRAHRSREKSYLFDVA 70

```

```

QY 326 TRLLQDSLGSAHSILIANIAPERRFYLDVTSALNFAARKEVIN 370
Db 302 TRLLQDSLGSAHSILIANIAPERRFYLDVTSALNFAARKEVIN 346

RESULT 7
US-10-425-115-320209
; Sequence 320209, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 320209
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_55097C.1.pep
; US-10-425-115-320209

Query Match      32.0%; Score 601; DB 17; Length 589;
Best Local Similarity 39.1%; Pred. No. 3.2e-50;
Matches 142; Conservative 72; Mismatches 113; Indels 36; Gaps 9;

QY 14 MAASAAISGAGRCRLSKIGATRRPPPARVRVAVRLRPFDGTAGASDPPCVRGMDSC- 72
Db 14 MAASAAISGAGRCRLSKIGATRRPPPARVRVAVRLRPFDGTAGASDPPCVRGMDSC- 72
QY 1 MAASTPARSSA---HLSQ-----PVRVLRVRPFLSSAASATAPCV-SLLGCH 46
Db 1 MAASTPARSSA---HLSQ-----PVRVLRVRPFLSSAASATAPCV-SLLGCH 46
QY 73 -----SLEIANWRNHOETLKQFDAPFGERSTQDDIYAGSVQPIRLHLLGONASVLA 127
Db 73 -----SLEIANWRNHOETLKQFDAPFGERSTQDDIYAGSVQPIRLHLLGONASVLA 127
QY 47 PGGGVTVQLKD-QHTSRSEQKLDAPFQEDSVSOIFDQEVRAVIPISFEGINATVAYG 105
Db 47 PGGGVTVQLKD-QHTSRSEQKLDAPFQEDSVSOIFDQEVRAVIPISFEGINATVAYG 105
QY 128 PTGACKTHMLGSPQPGVPIRALMDLQLTREAGAGRPWALSVTMSYLEIYQEKVLDL 187
Db 128 PTGACKTHMLGSPQPGVPIRALMDLQLTREAGAGRPWALSVTMSYLEIYQEKVLDL 187
QY 106 ATGSKTYTMOGTEDFPGLIPLAASTILALCT-----GTWC-SVEISYYEVTMERCYDL 158
Db 106 ATGSKTYTMOGTEDFPGLIPLAASTILALCT-----GTWC-SVEISYYEVTMERCYDL 158
QY 188 LDPASGDLVIREDCRGNILIPGLSOKPISSAFERHFLPASNRVTGATRLNQRSSRSH 247
Db 188 LDPASGDLVIREDCRGNILIPGLSOKPISSAFERHFLPASNRVTGATRLNQRSSRSH 247
QY 159 LEPKAKEIMALDDKGNMQLGSLVWPVRSMEEFQELYSIGVQRKVAHTGLNDVSSRSH 218
Db 159 LEPKAKEIMALDDKGNMQLGSLVWPVRSMEEFQELYSIGVQRKVAHTGLNDVSSRSH 218
QY 248 AVLLVKVDQERLAPFRQREGKLYLIDLAGSDNRRTGNKGLRLKESGAINTSFLVLGV 307
Db 248 AVLLVKVDQERLAPFRQREGKLYLIDLAGSDNRRTGNKGLRLKESGAINTSFLVLGV 307
QY 219 AVLRSVSTDVV-----KGKMLNLTSLXDLAGNEDNRRTCNIGIRLOESAKINQSLFALS 272
Db 219 AVLRSVSTDVV-----KGKMLNLTSLXDLAGNEDNRRTCNIGIRLOESAKINQSLFALS 272
QY 308 VDANQGLPRVPYRDSKLTRLQDSLGSAHSILIANIAPERRFYLDVTSALNFAARKE 367
Db 308 VDANQGLPRVPYRDSKLTRLQDSLGSAHSILIANIAPERRFYLDVTSALNFAARKE 367
QY 273 ISALNKNEHRIYRQSKLTRLLRDSLGSGSRVTVMIACLNPAB--YQESANTVSLAARSCH 330
Db 273 ISALNKNEHRIYRQSKLTRLLRDSLGSGSRVTVMIACLNPAB--YQESANTVSLAARSCH 330
QY 368 VIN 370
Db 368 VIN 370
QY 331 IEN 333
Db 331 IEN 333

RESULT 8
US-10-437-963-159031
; Sequence 159031, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad

```


APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 103943
LENGTH: 776
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_101324C.1.pap
US-10-437-963-103943

Query Match 29.0%; Score 544.5; DB 16; Length 776;
Best Local Similarity 35.3%; Pred. No. 1.9e-44; Indels 29; Gaps 6;
Matches 133; Conservative 71; Mismatches 144

QY 1 MAAGSTQORREMAAASAAISGAGRCRLSKIGATRRPPARVAVRLRPFVDGTAGA 60
DB LSLGGMAARLKTAGEAGNGDAAG-----SRIMVFLRPMRSKEKDA 222
QY 61 SDPPCVRGMDSCSLEIANWRNHQETLK-----YQDFYGERSTQDIYAGSVOPIL 112
DB GRSVCVTKVKKDVLTEFASSETDYLRLKVRGRHFCFDSFPDTTQAEVISTTSDLV 282
QY 113 RHLEGNASVLA YGPTGAGTHMLGSPQGVIPRALMDLLQLTREEGAGCRPWALS 172
DB EGVLRNGTVFCYATGAGTGYTMLGTWESPGVWVLAIKDLFTKVRQRSHDGNH---SI 339
QY 173 TMSLEYIYQEKVLDLDPASGDLVIREDCRGNILIPGLSQPISFPADFERHFLPASNR 232
DB QLSYLEVNETVRLDLSFGR-PLLRREDKQGTV-AAGLTHYRAYSTDEVKMLQOQGNR 397
QY 233 TVGATLNCRSSRSRAVLVKVDQRE-RLAPPROEGKLYLIDLAGSDNRRTGNKGLRL 291
DB TTEFTRVNTSRSRAILQVIEYRSIDGSGIVTRVGLSLIDLAGSERALATORTQS 457
QY 292 KESGAINISLFLVGLKVDALNQLPRVPYRSKLTRELQDSLGGSAAHSILIANIAPERRF 351
DB IEGANINRSLALSSCINALVEGKKHPIYRNSKLTQLLKDSLGGSCTNTMIANISPSNLS 517

QY 352 YLDTVSALNFAARSKEV 368
DB 518 FGETQNTLHWADRAKEI 534

RESULT 13
US-09-883-096-5
Sequence 5, Application US/09883096
Patent No. US20020110883A1
GENERAL INFORMATION:
APPLICANT: Beraud, Christophe
APPLICANT: Craven, Andrew
APPLICANT: Yu, Ming
APPLICANT: Sakowicz, Roman
APPLICANT: Patel, Umesh A.
APPLICANT: Davies, Katherine A.
TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE
FILE REFERENCE: 020552-001410US
CURRENT APPLICATION NUMBER: US/09/883,096
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 09/594,655
PRIOR FILING DATE: 2000-06-15
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 5
LENGTH: 383
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: HsKip3a
OTHER INFORMATION: fragment
OTHER INFORMATION: Amino acid sequence of HsKip3a fragment used in
OTHER INFORMATION: the ATPase assay (Figure 4).
US-09-883-096-5

Query Match 28.1%; Score 528; DB 9; Length 383;
Best Local Similarity 38.7%; Pred. No. 2.9e-43;
Matches 138; Conservative 46; Mismatches 127; Indels 46; Gaps 8;

QY 38 RPPPARVRVAVRLRPFV-----DG-----TAGASDPCCVGMDSCLSI 76
DB REPTPRELDSQR-RPVQVVDVRLVFNPEEPDGGFGLKGGTHDGGKKGKD----- 67
QY 77 ANWRNHQETLK YQDFADYGERSTQDIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGTHT 136
DB -----LTFVDFRVEGEAATQDDVFQHTHSLVDSFLOGYNCSVFAYGATGAGTHT 118
QY 137 MLGSPQGVIPRALMDLLQLTREEGAGCRPWALSVTMSLEYIYQEKVLDLDPASGDLV 196
DB MLGREGDPCGIM---YLTTVELYRLEARQEKHFEVLISYQEVYNEQIHDLEP-KGFLA 174
QY 197 IREDCRGNILIPGLSQPISFPADFERHFLPASNRRTVGATLNCRSSRSRAVLVKVDQ 256
DB IREDFDKGVVQGLSFHPQASAEQLLEILTRGNRRNTOHPTDANATSSRAIFQIFVKQ 234
QY 257 REELAPPRO-REGKLYLIDLAGSDNRRTGNKGLRLKESGAINISLFLVGLKVDALNQG 314
DB QRVPLGTQAVQVAKMSLIDLAGSERASTHAKGERLREGANINRSLALINVLNALADA 294
QY 315 LPR---VPYRDSKLTRELQDSLGGSAAHSILIANIAPERRFYLDTVSALNFAARSKEV 368
DB KGRKTHVPYRDSKLTRELKDSLGNCRVTWIAAISFSLTYEDTYNTLYADRAKEI 351

RESULT 14
US-09-883-096-2
Sequence 2, Application US/09883096
Patent No. US20020110883A1
GENERAL INFORMATION:
APPLICANT: Beraud, Christophe
APPLICANT: Craven, Andrew
APPLICANT: Yu, Ming
APPLICANT: Sakowicz, Roman
APPLICANT: Patel, Umesh A.
APPLICANT: Davies, Katherine A.
TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE
FILE REFERENCE: 020552-001410US
CURRENT APPLICATION NUMBER: US/09/883,096
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 09/594,655
PRIOR FILING DATE: 2000-06-15
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 864
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Amino acid sequence encoded by human kinesin motor
OTHER INFORMATION: protein gene HsKip3a (Figure 1).
OTHER INFORMATION: Description of Artificial Sequence: Amino acid
OTHER INFORMATION: sequence of HsKip3a.
US-09-883-096-2

Query Match 28.1%; Score 528; DB 9; Length 864;
Best Local Similarity 38.7%; Pred. No. 9.7e-43;
Matches 138; Conservative 46; Mismatches 127; Indels 46; Gaps 8;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:44:52 ; Search time 58.709 Seconds
(without alignments)
2081.726 Million cell updates/sec

Title: US-10-797-893-6

Perfect score: 1768

Sequence: 1 MGRCLSKIGATRRPPARV.....FYLDTSALNFAARSKEVIN 346

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1768	100.0	346	16 US-10-797-893-6	Sequence 6, Appli
2	1768	100.0	487	16 US-10-797-893-8	Sequence 8, Appli
3	1763	99.7	370	16 US-10-797-893-2	Sequence 2, Appli
4	1763	99.7	490	15 US-10-334-143-80	Sequence 80, Appli
5	1763	99.7	512	16 US-10-797-893-4	Sequence 4, Appli
6	1753	99.2	460	9 US-09-925-300-1228	Sequence 1228, Ap
7	597.5	33.8	377	16 US-10-437-963-159031	Sequence 159031,
8	593.5	33.6	589	17 US-10-425-115-320209	Sequence 320209,
9	559	31.6	548	15 US-10-108-260A-2692	Sequence 2692, Ap
10	550	31.1	783	17 US-10-425-115-332919	Sequence 332919,
11	550	31.1	854	15 US-10-425-114-59708	Sequence 59708, A
12	537	30.4	776	16 US-10-437-963-103943	Sequence 103943,
13	528	29.9	383	9 US-09-883-096-5	Sequence 5, Appli

14	528	29.9	864	9 US-09-883-096-2	Sequence 2, Appli
15	527	29.8	757	15 US-10-220-120-366	Sequence 366, App
16	527	29.8	757	16 US-10-363-829-316	Sequence 316, App
17	526	29.8	375	15 US-10-332-089-4	Sequence 4, Appli
18	526	29.8	409	15 US-10-332-089-6	Sequence 6, Appli
19	526	29.8	774	14 US-10-369-493-2361	Sequence 82, Appli
20	526	29.8	1388	14 US-10-146-473-82	Sequence 32, Appli
21	526	29.8	1388	15 US-10-173-999-32	Sequence 2, Appli
22	526	29.8	1388	15 US-10-332-089-2	Sequence 164, App
23	526	29.8	1388	15 US-10-188-832-164	Sequence 306, App
24	523.5	29.6	834	15 US-10-287-226-306	Sequence 142, App
25	521.5	29.5	1401	15 US-10-287-226-142	Sequence 1477, Ap
26	519.5	29.4	1826	16 US-10-408-765A-1477	Sequence 176714,
27	518.5	29.3	1382	16 US-10-437-963-176714	Sequence 302, App
28	518	29.3	1115	15 US-10-287-226-302	Sequence 670, App
29	518	29.3	1232	14 US-10-116-712-670	Sequence 2153, Ap
30	518	29.3	1232	16 US-10-408-765A-2153	Sequence 8, Appli
31	518	29.3	1235	15 US-10-334-143-8	Sequence 231231,
32	517.5	29.3	1001	17 US-10-425-115-231231	Sequence 314, App
33	517.5	29.3	1324	15 US-10-287-226-314	Sequence 7903, Ap
34	516	29.2	992	17 US-10-739-930-7903	Sequence 62748, A
35	515	29.1	1030	15 US-10-425-114-62748	Sequence 21903, A
36	513.5	29.0	805	14 US-10-369-493-21903	Sequence 33, Appli
37	513	29.0	1237	15 US-10-334-143-33	Sequence 1664, Ap
38	510	28.8	672	16 US-10-408-765A-1664	Sequence 1, Appli
39	509	28.8	1103	9 US-09-847-874A-1	Sequence 1, Appli
40	509	28.8	1103	14 US-10-458-162-1	Sequence 2, Appli
41	507	28.7	928	14 US-10-080-608A-23	Sequence 23, Appli
42	507	28.7	928	14 US-10-370-685-112	Sequence 112, App
43	507	28.7	1232	14 US-10-116-712-664	Sequence 664, App
44	507	28.7	1232	14 US-10-116-712-669	Sequence 669, App
45	504	28.5	338	9 US-09-883-096-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-10-797-893-6
; Sequence 6, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10797,893
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Human
; US-10-797-893-6

Query Match	100.0%;	Score 1768;	DB 16;	Length 346;
Best Local Similarity	100.0%;	Pred. No. 1.7e-168;		
Matches 346;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MGRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQE	60	
Db	1	MGRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQE	60	
QY	61	TLKQYDFAFGERSTQDDIYAGSVQPIRLHLLSGQNASVLAYGPTGAGKTHTMLGSPQ	120	
Db	61	TLKQYDFAFGERSTQDDIYAGSVQPIRLHLLSGQNASVLAYGPTGAGKTHTMLGSPQ	120	
QY	121	GVIPRALMDLLQLTREGAEGRPWALSVTMSYLEIYQEKVLDDLPASGDLVITREDCRGN	180	

```

Db 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180
QY 181 ILIPGLSQPISFADFERRHFLPASNRRTVGATRLNQRSSSRSHAVLLVKVDQERLAPFR 240
Db 181 ILIPGLSQPISFADFERRHFLPASNRRTVGATRLNQRSSSRSHAVLLVKVDQERLAPFR 240
QY 241 QREGKLYLIDLAGESENRRTGNKGLRKESGAINTSFLVLGKVDALNOGLPRVPYRDSK 300
Db 241 QREGKLYLIDLAGESENRRTGNKGLRKESGAINTSFLVLGKVDALNOGLPRVPYRDSK 300
QY 301 LTRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346
Db 301 LTRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346

```

RESULT 2

```

US-10-797-893-8
; Sequence 8, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/797,893
; PRIOR FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Human
US-10-797-893-8

```

```

Query Match 100.0%; Score 1768; DB 16; Length 487;
Best Local Similarity 100.0%; Pred. No. 2.8e-168;
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MGRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANRNHOE 60
Db 1 MGRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANRNHOE 60
QY 61 TLKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAQPTGAGKTHMLGSPEOP 120
Db 61 TLKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAQPTGAGKTHMLGSPEOP 120
QY 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180
Db 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180
QY 181 ILIPGLSQPISFADFERRHFLPASNRRTVGATRLNQRSSSRSHAVLLVKVDQERLAPFR 240
Db 181 ILIPGLSQPISFADFERRHFLPASNRRTVGATRLNQRSSSRSHAVLLVKVDQERLAPFR 240
QY 241 QREGKLYLIDLAGESENRRTGNKGLRKESGAINTSFLVLGKVDALNOGLPRVPYRDSK 300
Db 241 QREGKLYLIDLAGESENRRTGNKGLRKESGAINTSFLVLGKVDALNOGLPRVPYRDSK 300
QY 301 LTRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346
Db 301 LTRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346

```

RESULT 3

```

US-10-797-893-2
; Sequence 2, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/797,893
; PRIOR FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Human
US-10-797-893-2

```

```

Query Match 99.7%; Score 1763; DB 16; Length 370;
Best Local Similarity 100.0%; Pred. No. 5.9e-168;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANRNHOE 61
Db 26 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANRNHOE 85
QY 62 LKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAQPTGAGKTHMLGSPEOP 121
Db 86 LKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAQPTGAGKTHMLGSPEOP 145
QY 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 181
Db 146 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 205
QY 182 LIPGLSQPISFADFERRHFLPASNRRTVGATRLNQRSSSRSHAVLLVKVDQERLAPFR 241
Db 206 LIPGLSQPISFADFERRHFLPASNRRTVGATRLNQRSSSRSHAVLLVKVDQERLAPFR 265
QY 242 REGKLYLIDLAGESENRRTGNKGLRKESGAINTSFLVLGKVDALNOGLPRVPYRDSK 301
Db 266 REGKLYLIDLAGESENRRTGNKGLRKESGAINTSFLVLGKVDALNOGLPRVPYRDSK 325
QY 302 LTRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346
Db 326 LTRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 370

```

RESULT 4

```

US-10-334-143-80
; Sequence 80, Application US/10334143
; Publication No. US20040009549A1
; GENERAL INFORMATION:
; APPLICANT: GRIGORIIEV, IGOR VYACHESLAVOVICH
; APPLICANT: SUDARANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD
; FILE REFERENCE: 038602/1543
; CURRENT APPLICATION NUMBER: US/10/334,143
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: 60/343,169
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 80
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-334-143-80

```

```

Query Match 99.7%; Score 1763; DB 15; Length 490;
Best Local Similarity 100.0%; Pred. No. 9e-168;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 2 GRCRLSKIGATRRPPPARVRVAVRLRPFVDTAGASDPPPCVRGMDSCSLEIANWRNHQET 61
 Db 34 GRCRLSKIGATRRPPPARVRVAVRLRPFVDTAGASDPPPCVRGMDSCSLEIANWRNHQET 93
 QY 62 LKQFDPAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 121
 Db 94 LKQFDPAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 153
 QY 122 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 181
 Db 154 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 213
 QY 182 LIPGLSKQPISSPADFERHFLPASRNRVTGATRLNQRSSSHAVLLVKVDQRLAPFRQ 241
 Db 214 LIPGLSKQPISSPADFERHFLPASRNRVTGATRLNQRSSSHAVLLVKVDQRLAPFRQ 273
 QY 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVVDALNOGLPRVPYRDSKL 301
 Db 274 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVVDALNOGLPRVPYRDSKL 333
 QY 302 TRLLQDSLGGSAAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346
 Db 334 TRLLQDSLGGSAAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 378

RESULT 5
 US-10-797-893-4
 ; Sequence 4, Application US/10797893
 ; Publication No. US2004014237A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; TITLE OF INVENTION: Novel motor proteins and methods for
 ; FILE REFERENCE: 1044
 ; CURRENT APPLICATION NUMBER: US/10/797,893
 ; CURRENT FILING DATE: 2004-03-09
 ; PRIOR APPLICATION NUMBER: US/09/724,224
 ; PRIOR FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: 09/597,292
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 512
 ; TYPE: PRT
 ; ORGANISM: Human
 US-10-797-893-4

Query Match 99.7%; Score 1763; DB 16; Length 512;
 Best Local Similarity 100.0%; Pred. No. 9.6e-168;
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GRCRLSKIGATRRPPPARVRVAVRLRPFVDTAGASDPPPCVRGMDSCSLEIANWRNHQET 61
 Db 27 GRCRLSKIGATRRPPPARVRVAVRLRPFVDTAGASDPPPCVRGMDSCSLEIANWRNHQET 86
 QY 62 LKQFDPAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 121
 Db 87 LKQFDPAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 146
 QY 122 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 181
 Db 147 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 206
 QY 182 LIPGLSKQPISSPADFERHFLPASRNRVTGATRLNQRSSSHAVLLVKVDQRLAPFRQ 241
 Db 207 LIPGLSKQPISSPADFERHFLPASRNRVTGATRLNQRSSSHAVLLVKVDQRLAPFRQ 266
 QY 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVVDALNOGLPRVPYRDSKL 301
 Db 267 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVVDALNOGLPRVPYRDSKL 326
 QY 302 TRLLQDSLGGSAAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346

Db 327 TRLLQDSLGGSAAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 371
 RESULT 6
 US-09-925-300-1228
 ; Sequence 1228, Application US/09925300
 ; Patent No. US20020151681A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Craig Rosen,
 ; APPLICANT: Steve Ruben
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA101
 ; CURRENT APPLICATION NUMBER: US/09/925,300
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05988
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 1890
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1228
 ; LENGTH: 460
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (75)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (147)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (435)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 US-09-925-300-1228

Query Match 99.2%; Score 1753; DB 9; Length 460;
 Best Local Similarity 99.4%; Pred. No. 8.2e-167;
 Matches 343; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 GRCRLSKIGATRRPPPARVRVAVRLRPFVDTAGASDPPPCVRGMDSCSLEIANWRNHQET 61
 Db 4 GRCRLSKIGATRRPPPARVRVAVRLRPFVDTAGASDPPPCVRGMDSCSLEIANWRNHQET 63
 QY 62 LKQFDPAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 121
 Db 64 LKQFDPAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 123
 QY 122 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 181
 Db 124 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 183
 QY 182 LIPGLSKQPISSPADFERHFLPASRNRVTGATRLNQRSSSHAVLLVKVDQRLAPFRQ 241
 Db 184 LIPGLSKQPISSPADFERHFLPASRNRVTGATRLNQRSSSHAVLLVKVDQRLAPFRQ 243
 QY 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVVDALNOGLPRVPYRDSKL 301
 Db 244 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVVDALNOGLPRVPYRDSKL 303
 QY 302 TRLLQDSLGGSAAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346
 Db 304 TRLLQDSLGGSAAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 348

RESULT 7
 US-10-437-963-159031
 ; Sequence 159031, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.

```

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 159031
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(377)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_58448C.1.psp
; US-10-437-963-159031

Query Match      33.8%; Score 597.5; DB 16; Length 377;
Best Local Similarity 39.7%; Pred. No. 6.5e-51;
Matches 138; Conservative 69; Mismatches 112; Indels 29; Gaps 8;

QY      6 LSKIGATRRPPARVRVRLRPFDGTAGADPPCVRGMDSCSLEIANRNHQTLLK- 63
DB      1 MATAAATQSP---VRVRLVRPHLPSEANSARACVGLGSHPGGVTVQLKDQYTSN 57

QY      64 --YQDFAFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAIGPTGAGKTHMLGSPGPG 121
DB      58 ECVKLDAPFGQSSRVCEIFDQESAVIPGIFEGTNATFAYGATGSGKTYTMOGTEDLPG 117

QY      122 VTPRALMDLLQTLREBAGRPWALSVMYSLEYIQEKVLDLDPASGLDVIREDCRNI 181
DB      118 LMXSTVLALCTGT-----WC-SVEISYEVYMERCYDLEFPKAKEIMALDDKGNMQ 167

QY      182 LIPGLSOKPISSFADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQRLAPFRQ 241
DB      168 QLKGLAWVPVRSLEEFHEIYSIGVQRKVATHTGLNDVSSRSHAVLSIRITT-----DV 220

QY      242 REGKLYLI---DLAGEDNRRRTGNKGLRKESGAINTSFLVLGKVVDALNQGLPRVPYRD 298
DB      221 VKGKMLITSXDLAGNEDNRRTCNEGIRLQESAKINQSLFALSNNVISALNKKEPRIPYRE 280

QY      299 SKLTRILODSLGSASHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346
DB      281 SKLTRILODSLGSASHVMIACLNPVE--YQEAHTVTSLAARSHTVN 326

RESULT 8
US-10-425-115-320209
; Sequence 320209, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 320209
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:

```

```

; OTHER INFORMATION: Clone ID: MRT4577_55097C.1.psp
; US-10-425-115-320209

Query Match      33.6%; Score 593.5; DB 17; Length 589;
Best Local Similarity 39.2%; Pred. No. 3.2e-50;
Matches 135; Conservative 71; Mismatches 107; Indels 31; Gaps 8;

QY      17 PAR-----VRVAVRLRPFDGTAGASDPPCVRGMDSC-----SLEIANRNHQTLL 62
DB      7 PARSSAHLSPQVRVRLVRPFLSSEASATAPCV-SLLGCHPGGVTVQKX-QHTSRSE 64

QY      63 KYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAIGPTGAGKTHMLGSPGPG 122
DB      65 QYKLDAPFRQEDSVSQIFDQEVRAVIPSIFEGINATVFAYGATGSGKTYTMOGTEDFPL 124

QY      123 IPRALMDLLQTLREBAGRPWALSVMYSLEYIQEKVLDLDPASGLDVIREDCRNI 182
DB      125 IPLAASTILALCT-----GTWC-SVEISYEVYMERCYDLEFPKAKEIMALDDKGNMQ 177

QY      183 IPGLSOKPISSFADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQRLAPFRQ 242
DB      178 LKGLSWVPVRSMEEFQELISIGVQRKVATHTGLNDVSSRSHAVLSRVSSTDVV----- 231

QY      243 EGKLYLIDLAGSEDNRRRTGNKGLRKESGAINTSFLVLGKVVDALNQGLPRVPYRD 302
DB      232 KGKLNLDLAGSEDNRRRTGNKGLRKESGAINTSFLVLGKVVDALNQGLPRVPYRD 291

QY      303 RLLODSLGSASHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346
DB      292 RLRLDSLGSASHSILIANIAPERRFYLDTVSALNFAARSKEVIN 333

RESULT 9
US-10-108-260A-2692
; Sequence 2692, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1 full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2692
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-108-260A-2692

Query Match      31.6%; Score 559; DB 15; Length 548;
Best Local Similarity 38.9%; Pred. No. 8.3e-47;
Matches 132; Conservative 70; Mismatches 121; Indels 16; Gaps 7;

QY      19 RRVAVRLRPF--VDGTAGAS-----DPPCVRGMDSCSLEIANRNHQTLLK-YQDFAF 69
DB      11 QLMVALRVRLPISVAELEBEGATLIAHKVDEQVWLMDPMDPDITLRAHRSREKSYLFDVA 70

QY      70 YGERSTQODIYAGSVQPIRLHLLLEGONASVLAIGPTGAGKTHMLGSPGPGVPRALMD 129
DB      71 FDFATQEMVYQATTKSLIEGVISGYNATVFAYGATGCGKTYTMOGTEDFPIVQTLND 130

QY      130 LLOTLREBAGRPWALSVMYSLEYIQEKVLDLDPASGLDVIREDCRNI 189
DB      131 LFRATEETSND---MEYEVMSLEYIYNEMRLDNLPSGLYLELREDSKGVIVAGITEV 187

QY      190 FISSPADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQRLAPFRQ--REGKLY 247
DB      188 STINAKEIMQLMKGNRQRTQPTAANTSSRSHAVLQVTVRQSRVKNILQEVQRGLF 247

QY      248 LIDLAGSEDNRRRTGNKGLRKESGAINTSFLVLGKVVDAL--NOGLPR--VPYRDSKLTLL 305
DB      248 MIDLAGSERASQTQNRGQMKEGAHINRSLALGNCINALSDKGNKYNINRDSKLTLL 307

```


QY 18 ARVRVAVLRPFVDTAGASDPPCVRGMDSCSLEIANWRNHQETLK-----YQDAP 69
 Db 204 SRIMFVRLRMSRKEKDGSRSCVKIYNKDVLTTEFASSETDYLRKVRGRHFCFDS 263
 QY 70 YGERSTQODIVAGSVQPIRLHLEGONASVLAYGPTGAGKTHMLGSPQPGVPRALMD 129
 Db 264 PDDTTQAEVSTTSDIVEGVQGRNTVCYGNATGAGKYTMLGTWESFGVWLAKD 323
 QY 130 LQLTREBEGAGRPWALSVTWMSYLEIYQEKVLDLDPASGLVIREDCRGNILIPGLSQK 189
 Db 324 LFTKVRQRSHDGNH---SIQLSYLEVNETVRDLSPGR-PLLLREDKQGTV-AAGLTHY 378
 QY 190 PISSFADFERHPLASRNTVTGATRLNQRSSRSHAVLLVKVDOR-RLAPFRQEGKLYL 248
 Db 379 RAYSTDVEMKLIQOQNRTTEPRVNETSSRSHAILQVIVEYRSDGGSIVTRVGKLSL 438
 QY 249 IDLAGSDNRRTGNKGLRKESGAINTSFLVLGVVDALNOGLPRVPYRDSKLTLLQDS 308
 Db 439 IDLAGSERALATDQRTQRTSIEGANINRSLLALSSCINALVEGKKHPIYRNSKLTQLLKDS 498
 QY 309 LGGSAHSILIANIAPERRFYLDVTVSALNFAARSKEV 344
 Db 499 LGGSCNTVMIANISPSNLSFGSTQNTLHWADRAKEI 534

RESULT 13
 US-09-883-096-5
 ; Sequence 5, Application US/09883096
 ; Patent No. US20020110883A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Craven, Andrew
 ; APPLICANT: Yu, Ming
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Patel, Umesh A.
 ; APPLICANT: Davies, Katherine A.
 ; TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE
 ; FILE REFERENCE: 020552-001410US
 ; CURRENT APPLICATION NUMBER: US/09/883,096
 ; CURRENT FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: US 09/594,655
 ; PRIOR FILING DATE: 2000-06-15
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 5
 ; LENGTH: 383
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Hskip3a
 ; OTHER INFORMATION: fragment
 ; OTHER INFORMATION: Amino acid sequence of Hskip3a fragment used in
 ; OTHER INFORMATION: the ATPase assay (Figure 4).
 US-09-883-096-5

Query Match 29.9%; Score 528; DB 9; Length 383;
 Best Local Similarity 38.7%; Pred. No. 6.3e-44;
 Matches 138; Conservative 46; Mismatches 127; Indels 46; Gaps 8;
 QY 14 RPPPARVRVAVLRPFV-----DG-----TAGASDPPCVRGMDSCSLEI 52
 Db 15 RPPTRELDQSQR-RPVVQVVDRLVFNPEPDGGFGLKGGTHDGPKKGKD----- 67
 QY 53 ANWRNHQETLVKQFDAFYGERSTQODIYAGSVQPIRLHLEGONASVLAYGPTGAGKTH 112
 Db 68 -----LTFVDRVFGEAATQDDVFQHTTHSVLDSPLOGYNCSVFAYGATGAGKTH 118
 QY 113 MLGSPQPGVPRALMDLLQLTREBEGAGRPWALSVTWMSYLEIYQEKVLDLDPASGLV 172
 Db 119 MLGREGDPMGIM---YLTVELYRLRLEARQOEKHEVLISYQEVYNEQIHDLEP-KGFLA 174
 QY 173 IREDCRGNILIPGLSQKPISSFADFERHPLASRNTVTGATRLNQRSSRSHAVLLVKVDQ 232
 Db 175 IREDDPKGVVQGLSFHQPASAEQLLELTGTGNRRNTQHTPDANATSSRSHAIQFIVKQ 234
 QY 233 RERLAPFRQ---REGKLYLDLAGSDNRRTGNKGLRKESGAINTSFLVLGVVDALNOG 290
 Db 235 QDRVPGLTQAVQAKMSLIDLAGSERASSHTAKGERLREGANINRSLLALINVLNADADA 294
 QY 291 LPR---VPYRDSKLTLLQDSLGGSAHSILIANIAPERRFYLDVTVSALNFAARSKEV 344
 Db 295 KGRKTHVPYRDSKLTLLKDSLGGNCRTVMIAAISPSLSLTEDTYNTLYADRAKEI 351
 QY 173 IREDCRGNILIPGLSQKPISSFADFERHPLASRNTVTGATRLNQRSSRSHAVLLVKVDQ 232

Db 175 IREDDPKGVVQGLSFHQPASAEQLLELTGTGNRRNTQHTPDANATSSRSHAIQFIVKQ 234
 QY 233 RERLAPFRQ---REGKLYLDLAGSDNRRTGNKGLRKESGAINTSFLVLGVVDALNOG 290
 Db 235 QDRVPGLTQAVQAKMSLIDLAGSERASSHTAKGERLREGANINRSLLALINVLNADADA 294
 QY 291 LPR---VPYRDSKLTLLQDSLGGSAHSILIANIAPERRFYLDVTVSALNFAARSKEV 344
 Db 295 KGRKTHVPYRDSKLTLLKDSLGGNCRTVMIAAISPSLSLTEDTYNTLYADRAKEI 351

RESULT 14
 US-09-883-096-2
 ; Sequence 2, Application US/09883096
 ; Patent No. US20020110883A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Craven, Andrew
 ; APPLICANT: Yu, Ming
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Patel, Umesh A.
 ; APPLICANT: Davies, Katherine A.
 ; TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE
 ; FILE REFERENCE: 020552-001410US
 ; CURRENT APPLICATION NUMBER: US/09/883,096
 ; CURRENT FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: US 09/594,655
 ; PRIOR FILING DATE: 2000-06-15
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 864
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Amino acid sequence encoded by human kinesin motor
 ; OTHER INFORMATION: protein gene Hskip3a (Figure 1).
 ; OTHER INFORMATION: Description of Artificial Sequence: Amino acid
 ; OTHER INFORMATION: sequence of Hskip3a.
 US-09-883-096-2

Query Match 29.9%; Score 528; DB 9; Length 864;
 Best Local Similarity 38.7%; Pred. No. 2.1e-43;
 Matches 138; Conservative 46; Mismatches 127; Indels 46; Gaps 8;
 QY 14 RPPPARVRVAVLRPFV-----DG-----TAGASDPPCVRGMDSCSLEI 52
 Db 15 RPPTRELDQSQR-RPVVQVVDRLVFNPEPDGGFGLKGGTHDGPKKGKD----- 67
 QY 53 ANWRNHQETLVKQFDAFYGERSTQODIYAGSVQPIRLHLEGONASVLAYGPTGAGKTH 112
 Db 68 -----LTFVDRVFGEAATQDDVFQHTTHSVLDSPLOGYNCSVFAYGATGAGKTH 118
 QY 113 MLGSPQPGVPRALMDLLQLTREBEGAGRPWALSVTWMSYLEIYQEKVLDLDPASGLV 172
 Db 119 MLGREGDPMGIM---YLTVELYRLRLEARQOEKHEVLISYQEVYNEQIHDLEP-KGFLA 174
 QY 173 IREDCRGNILIPGLSQKPISSFADFERHPLASRNTVTGATRLNQRSSRSHAVLLVKVDQ 232
 Db 175 IREDDPKGVVQGLSFHQPASAEQLLELTGTGNRRNTQHTPDANATSSRSHAIQFIVKQ 234
 QY 233 RERLAPFRQ---REGKLYLDLAGSDNRRTGNKGLRKESGAINTSFLVLGVVDALNOG 290
 Db 235 QDRVPGLTQAVQAKMSLIDLAGSERASSHTAKGERLREGANINRSLLALINVLNADADA 294
 QY 291 LPR---VPYRDSKLTLLQDSLGGSAHSILIANIAPERRFYLDVTVSALNFAARSKEV 344
 Db 295 KGRKTHVPYRDSKLTLLKDSLGGNCRTVMIAAISPSLSLTEDTYNTLYADRAKEI 351

RESULT 15
 US-10-220-120-366

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 ; Search time 13.5172 Seconds
(without alignments)
1697.543 Million cell updates/sec

Title: US-10-797-893-6
Perfect score: 1768
Sequence: 1 MGRCLSKIGATRRPPPARV.....FYLDTSALNFAARKEVIN 346

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pcp.*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB.pcp.*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB.pcp.*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB.pcp.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pcp.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1768	100.0	346	3	US-09-724-224-6
2	1768	100.0	346	4	US-10-093-317-6
3	1768	100.0	487	3	US-09-724-224-8
4	1768	100.0	487	4	US-10-093-317-8
5	1763	99.7	370	3	US-09-724-224-2
6	1763	99.7	370	4	US-10-093-317-2
7	1763	99.7	512	3	US-09-724-224-4
8	1763	99.7	512	4	US-10-093-317-4
9	1763	99.7	665	3	US-09-595-684B-35
10	549	31.1	355	3	US-09-724-511-4
11	549	31.1	355	4	US-09-723-097-4
12	549	31.1	355	4	US-09-632-344-4
13	549	31.1	367	3	US-09-724-511-2
14	549	31.1	367	4	US-09-723-097-2
15	549	31.1	367	4	US-09-632-344-2
16	528	29.9	383	4	US-09-883-096-5
17	528	29.9	864	4	US-09-883-096-2
18	526	29.8	375	3	US-09-572-191-4
19	526	29.8	375	3	US-09-723-262-4
20	526	29.8	375	3	US-09-723-219-4
21	526	29.8	409	3	US-09-572-191-6
22	526	29.8	409	3	US-09-723-262-6
23	526	29.8	409	3	US-09-723-219-6
24	526	29.8	1388	3	US-09-572-191-2
25	526	29.8	1388	3	US-09-723-262-2
26	526	29.8	1388	3	US-09-723-219-2
27	521.5	29.5	522	4	US-09-592-054-4

28	521.5	29.5	1279	3	US-09-724-517-2	Sequence 2, Appli
29	521.5	29.5	1279	4	US-09-641-807A-2	Sequence 2, Appli
30	521.5	29.5	1279	4	US-09-723-096-2	Sequence 2, Appli
31	518.5	29.3	341	3	US-09-724-517-4	Sequence 4, Appli
32	518.5	29.3	341	4	US-09-641-807A-4	Sequence 4, Appli
33	518.5	29.3	341	4	US-09-723-096-4	Sequence 4, Appli
34	518	29.3	473	4	US-09-592-054-6	Sequence 6, Appli
35	517.5	29.3	342	3	US-09-641-806-2	Sequence 2, Appli
36	517.5	29.3	342	4	US-09-723-129-2	Sequence 2, Appli
37	517.5	29.3	342	4	US-09-722-862-2	Sequence 2, Appli
38	514.5	29.1	337	3	US-09-641-806-4	Sequence 4, Appli
39	514.5	29.1	337	4	US-09-723-129-4	Sequence 4, Appli
40	514.5	29.1	337	4	US-09-722-862-4	Sequence 4, Appli
41	514	29.1	1066	3	US-09-541-782-8	Sequence 8, Appli
42	514	29.1	1066	4	US-09-723-820-8	Sequence 8, Appli
43	514	29.1	1066	4	US-10-270-085-8	Sequence 8, Appli
44	514	29.1	1690	4	US-09-595-684B-39	Sequence 39, Appli
45	511.5	28.9	299	3	US-09-621-233-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-724-224-6
; Sequence 6, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
; FILE REFERENCE: their use
; CURRENT APPLICATION NUMBER: US/09/724,224
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Human
US-09-724-224-6

Query Match	100.0%	Score	1768	DB	3	Length	346
Best Local Similarity	100.0%	Pred. No.	1.2e-180				
Matches	346	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MGRCLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRHQE	60				
Db	1	MGRCLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRHQE	60				
Qy	61	TLKQFDAPFYGERSTQDIYAGSVQPIRLHLLRHLLEGONASVLAAYGPTGAGKTHTMLGSPEQP	120				
Db	61	TLKQFDAPFYGERSTQDIYAGSVQPIRLHLLRHLLEGONASVLAAYGPTGAGKTHTMLGSPEQP	120				
Qy	121	GVTPRALMDLLOLTREGAGRPWALSVTMSYLEIQEKVLDLLDPASGDLVREDCRGN	180				
Db	121	GVTPRALMDLLOLTREGAGRPWALSVTMSYLEIQEKVLDLLDPASGDLVREDCRGN	180				
Qy	181	ILIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR	240				
Db	181	ILIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR	240				
Qy	241	QREGKLYLIDLAGSDNRRTGNKGLRKSAGINTSLFVLGKVVADALNQLPRVPYRDSK	300				
Db	241	QREGKLYLIDLAGSDNRRTGNKGLRKSAGINTSLFVLGKVVADALNQLPRVPYRDSK	300				
Qy	301	LTELLODSLGGSAHSILIANIAPRRFYLDTSALNFAARKEVIN	346				
Db	301	LTELLODSLGGSAHSILIANIAPRRFYLDTSALNFAARKEVIN	346				

Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MGRCLSKIGATRRPPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
DB	1 MGRCLSKIGATRRPPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
QY	61 TLKYQDFAFYGERSTOODIYAGSVQPIRLHLEGGQNASVLAQPTGAGKTHTMLGSPQ 120
DB	61 TLKYQDFAFYGERSTOODIYAGSVQPIRLHLEGGQNASVLAQPTGAGKTHTMLGSPQ 120
QY	121 GVIPRALMDLLQLTREAGRGPWALSVMYSLEYIQEKVLDLDPASGDLVIREDCRG 180
DB	121 GVIPRALMDLLQLTREAGRGPWALSVMYSLEYIQEKVLDLDPASGDLVIREDCRG 180
QY	181 ILIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPR 240
DB	181 ILIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPR 240
QY	241 QREGKLYLIDLAGESENRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVYRDSK 300
DB	241 QREGKLYLIDLAGESENRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVYRDSK 300
QY	301 LTRLQDLSGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
DB	301 LTRLQDLSGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
RESULT 4	
US-10-093-317-8	
; Sequence 8, Application US/10093317	
; Patent No. 6762043	
; GENERAL INFORMATION:	
; APPLICANT: Beraud, Christophe	
; TITLE OF INVENTION: their use	
; FILE REFERENCE: 1044	
; CURRENT APPLICATION NUMBER: US/10/093,317	
; PRIOR FILING DATE: 2002-03-06	
; PRIOR APPLICATION NUMBER: 09/724,224	
; PRIOR FILING DATE: 2000-11-28	
; NUMBER OF SEQ ID NOS: 8	
; SOFTWARE: FastSeq for Windows Version 4.0	
; SEQ ID NO 6	
; LENGTH: 346	
; TYPE: PRT	
; ORGANISM: Human	
US-10-093-317-6	
Query Match 100.0%; Score 1768; DB 4; Length 346;	
Best Local Similarity 100.0%; Pred. No. 1.2e-180;	
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MGRCLSKIGATRRPPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
DB	1 MGRCLSKIGATRRPPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
QY	61 TLKYQDFAFYGERSTOODIYAGSVQPIRLHLEGGQNASVLAQPTGAGKTHTMLGSPQ 120
DB	61 TLKYQDFAFYGERSTOODIYAGSVQPIRLHLEGGQNASVLAQPTGAGKTHTMLGSPQ 120
QY	121 GVIPRALMDLLQLTREAGRGPWALSVMYSLEYIQEKVLDLDPASGDLVIREDCRG 180
DB	121 GVIPRALMDLLQLTREAGRGPWALSVMYSLEYIQEKVLDLDPASGDLVIREDCRG 180
QY	181 ILIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPR 240
DB	181 ILIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPR 240
QY	241 QREGKLYLIDLAGESENRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVYRDSK 300
DB	241 QREGKLYLIDLAGESENRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVYRDSK 300
QY	301 LTRLQDLSGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
DB	301 LTRLQDLSGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
RESULT 3	
US-09-724-224-8	
; Sequence 8, Application US/09724224	
; Patent No. 6387644	
; GENERAL INFORMATION:	
; APPLICANT: Beraud, Christophe	
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for	
; FILE REFERENCE: 1044	
; CURRENT APPLICATION NUMBER: US/09/724,224	
; PRIOR FILING DATE: 2000-11-28	
; PRIOR APPLICATION NUMBER: 09/597,292	
; PRIOR FILING DATE: 2000-06-20	
; NUMBER OF SEQ ID NOS: 8	
; SOFTWARE: FastSeq for Windows Version 4.0	
; SEQ ID NO 8	
; LENGTH: 487	
; TYPE: PRT	
; ORGANISM: Human	
US-09-724-224-8	
Query Match 100.0%; Score 1768; DB 3; Length 487;	
Best Local Similarity 100.0%; Pred. No. 2.1e-180;	
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MGRCLSKIGATRRPPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
DB	1 MGRCLSKIGATRRPPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
QY	61 TLKYQDFAFYGERSTOODIYAGSVQPIRLHLEGGQNASVLAQPTGAGKTHTMLGSPQ 120
DB	61 TLKYQDFAFYGERSTOODIYAGSVQPIRLHLEGGQNASVLAQPTGAGKTHTMLGSPQ 120
QY	121 GVIPRALMDLLQLTREAGRGPWALSVMYSLEYIQEKVLDLDPASGDLVIREDCRG 180
DB	121 GVIPRALMDLLQLTREAGRGPWALSVMYSLEYIQEKVLDLDPASGDLVIREDCRG 180
QY	181 ILIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPR 240
DB	181 ILIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPR 240
QY	241 QREGKLYLIDLAGESENRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVYRDSK 300
DB	241 QREGKLYLIDLAGESENRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVYRDSK 300
QY	301 LTRLQDLSGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
DB	301 LTRLQDLSGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
Query Match 100.0%; Score 1768; DB 3; Length 487;	
Best Local Similarity 100.0%; Pred. No. 2.1e-180;	
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MGRCLSKIGATRRPPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
DB	1 MGRCLSKIGATRRPPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
QY	61 TLKYQDFAFYGERSTOODIYAGSVQPIRLHLEGGQNASVLAQPTGAGKTHTMLGSPQ 120
DB	61 TLKYQDFAFYGERSTOODIYAGSVQPIRLHLEGGQNASVLAQPTGAGKTHTMLGSPQ 120
QY	121 GVIPRALMDLLQLTREAGRGPWALSVMYSLEYIQEKVLDLDPASGDLVIREDCRG 180
DB	121 GVIPRALMDLLQLTREAGRGPWALSVMYSLEYIQEKVLDLDPASGDLVIREDCRG 180
QY	181 ILIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPR 240
DB	181 ILIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPR 240
QY	241 QREGKLYLIDLAGESENRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVYRDSK 300
DB	241 QREGKLYLIDLAGESENRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVYRDSK 300
QY	301 LTRLQDLSGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
DB	301 LTRLQDLSGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

```
|||||
Db 301 LTRLLQSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVIN 346

RESULT 5
US-09-724-224-2
; Sequence 2, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Human
US-09-724-224-2

Query Match
Best Local Similarity 99.7%; Score 1763; DB 3; Length 370;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRCLSKIGATRRPPPARVRVAVRLRPFVDTAGASDPPCCVGRGMDSCSLEIANWRNHQET 61
Db 26 GRCLSKIGATRRPPPARVRVAVRLRPFVDTAGASDPPCCVGRGMDSCSLEIANWRNHQET 85

Qy 62 LKQFDIFYGERSTQDDIYAGSVQPIRLHLEGNASVLAAYGPTGAGKTHTMLGSPQPG 121
Db 86 LKQFDIFYGERSTQDDIYAGSVQPIRLHLEGNASVLAAYGPTGAGKTHTMLGSPQPG 145

Qy 122 VIPRALMDLLQLTREAGRPAWLSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 181
Db 146 VIPRALMDLLQLTREAGRPAWLSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 205

Qy 182 LIPCLSKQPISSFADFERHFLPASRNTVGATRLNQSRSSHAVLLVKVDQRLAPPRQ 241
Db 206 LIPCLSKQPISSFADFERHFLPASRNTVGATRLNQSRSSHAVLLVKVDQRLAPPRQ 265

Qy 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNQGLPRVPYRDSKL 301
Db 266 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNQGLPRVPYRDSKL 325

Qy 302 TRLLQDSLGSASHILIANIAPERRFYLDTVSALNFAARSKVIN 346
Db 326 TRLLQDSLGSASHILIANIAPERRFYLDTVSALNFAARSKVIN 370

RESULT 7
US-09-724-224-4
; Sequence 4, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Human
US-09-724-224-4

Query Match
Best Local Similarity 99.7%; Score 1763; DB 3; Length 512;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRCLSKIGATRRPPPARVRVAVRLRPFVDTAGASDPPCCVGRGMDSCSLEIANWRNHQET 61
Db 27 GRCLSKIGATRRPPPARVRVAVRLRPFVDTAGASDPPCCVGRGMDSCSLEIANWRNHQET 86

Qy 62 LKQFDIFYGERSTQDDIYAGSVQPIRLHLEGNASVLAAYGPTGAGKTHTMLGSPQPG 121
Db 87 LKQFDIFYGERSTQDDIYAGSVQPIRLHLEGNASVLAAYGPTGAGKTHTMLGSPQPG 146

Qy 122 VIPRALMDLLQLTREAGRPAWLSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 181
Db 147 VIPRALMDLLQLTREAGRPAWLSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 206

Qy 182 LIPCLSKQPISSFADFERHFLPASRNTVGATRLNQSRSSHAVLLVKVDQRLAPPRQ 241
Db 207 LIPCLSKQPISSFADFERHFLPASRNTVGATRLNQSRSSHAVLLVKVDQRLAPPRQ 266

Qy 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNQGLPRVPYRDSKL 301

|||||
Db 301 LTRLLQSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVIN 346

RESULT 6
US-10-093-317-2
; Sequence 2, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093,317
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Human
```

```

Db 267 REGKYLIDLAGSDNRRTGNKGLRKGSGAINTSLFVLGKVDALNQLPRVPYRDSKL 326
Qy 302 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
Db 327 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 371

RESULT 8
US-10-093-317-4
; Sequence 4, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093,317
; PRIOR FILING DATE: 2002-03-06
; PRIOR FILING DATE: 2002-03-06
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Human
US-10-093-317-4
Query Match 99.7%; Score 1763; DB 4; Length 512;
Best Local Similarity 100.0%; Pred. No. 7.6e-180;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGSDPPCVRGMDSCSLEIANWRNHQET 61
Db 27 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGSDPPCVRGMDSCSLEIANWRNHQET 86
Qy 62 LKQFDIFYGERSTQDDIYAGSVQPIRLHLEGNASVLAGYPTGAGKTHMLGSPQPG 121
Db 87 LKQFDIFYGERSTQDDIYAGSVQPIRLHLEGNASVLAGYPTGAGKTHMLGSPQPG 146
Qy 122 VIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIQEKVLDDLPASGDLVIREDCRNI 181
Db 147 VIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIQEKVLDDLPASGDLVIREDCRNI 206
Qy 182 LIPGLSKQPISSFADFERHFLPASRNTVGTATRLNQRSSSRSHAVLLVKVDQERLAPFRQ 241
Db 147 VIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIQEKVLDDLPASGDLVIREDCRNI 206
Qy 182 LIPGLSKQPISSFADFERHFLPASRNTVGTATRLNQRSSSRSHAVLLVKVDQERLAPFRQ 241
Db 207 LIPGLSKQPISSFADFERHFLPASRNTVGTATRLNQRSSSRSHAVLLVKVDQERLAPFRQ 266
Qy 242 REGKYLIDLAGSDNRRTGNKGLRKGSGAINTSLFVLGKVDALNQLPRVPYRDSKL 301
Db 267 REGKYLIDLAGSDNRRTGNKGLRKGSGAINTSLFVLGKVDALNQLPRVPYRDSKL 326
Qy 302 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
Db 327 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 371

```

```

RESULT 9
US-09-595-684B-35
; Sequence 35, Application US/09595684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Ohashi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; FILE REFERENCE: cytop036
; CURRENT APPLICATION NUMBER: US/09/595,684B

Qy 267 REGKYLIDLAGSDNRRTGNKGLRKGSGAINTSLFVLGKVDALNQLPRVPYRDSKL 326
Db 302 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
Qy 302 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
Db 327 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 371

RESULT 10
US-09-724-511-4
; Sequence 4, Application US/09724511
; Patent No. 6391601
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6391601 motor proteins and methods for
; FILE REFERENCE: 1046
; CURRENT APPLICATION NUMBER: US/09/724,511
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Human
US-09-724-511-4
Query Match 31.1%; Score 549; DB 3; Length 355;
Best Local Similarity 39.0%; Pred. No. 4.6e-50;
Matches 141; Conservative 59; Mismatches 114; Indels 48; Gaps 11;

Qy 20 VRVAVLRP-----FVDGTAGSDP-----PCVRGMDSCSLEIANWRNHQ 59
Db 4 MKYVVRVREPTKEKAAAGFHKKVHVVDKHLVDFDPKQEVSTFHGKKTNNQVKKQN-- 61
Qy 60 ETLKYQFDIFYGERSTQDDIYAGSVQPIRLHLEGNASVLAGYPTGAGKTHMLGSPQ 119
Db 62 KDLKFVDFADVDETSTQSEVFETTKPILRSFLNGYNTVCTVLAAYGATGAGKTHMLGSADE 121

```



```
QY 120 PGVPRALMDLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDPASGDLVIREDCRG 179
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 PGVWYLTMLHLKCMDEIKEE---ICSTAVSYLEVYNEQIRDLL-VNSGPLAVREDTQK 177

QY 180 NILIPGLS-QKPISSPADFERHFLP-ASRNRVTGATRLNQRSSRSRSHAVLLVKVDQERLA 237
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 178 GVVVHGLTLHQPSS--EETLHLLDNGKNRTOHPTDMNATSSRSRSHAVFQIYLRQODKTA 235

QY 238 PFRQ--REGKLYLIDLAGEEDNRRTGNKGLRKESCAINTSLFVLGKVVDAL----- 287
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 236 SINQNVRIAKMSLIDLAGESTSGAKGTRFVEGTNINRSLLALGNVINALADSKISFL 295

QY 288 -----NQGLPRVPYRDSKLTLLQDSLGGSAAHSILIANIAPERFYLDTVSALNFAARSK 342
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 296 CFQRKNQ---HIPYRNSKLTLLKDSLGNCQTIMIAAVSPSSVFYDDTYNTLKYANRAK 352

QY 343 EV 344
||: : : : :
Db 353 DI 354

RESULT 11
US-09-723-097-4
; Sequence 4, Application US/09723097
; Patent No. 6492151
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6492151el motor proteins and methods for
; FILE REFERENCE: 1046
; CURRENT APPLICATION NUMBER: US/09/723,097
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/632,344
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Human
US-09-723-097-4

Query Match 31.1%; Score 549; DB 4; Length 355;
Best Local Similarity 39.0%; Pred. No. 4.6e-50;
Matches 141; Conservative 59; Mismatches 114; Indels 48; Gaps 11;

QY 20 VRVAVRLRP-----FVDGTAGASDP-----PCVRGMDSCSLEIANWRNHQ 59
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 MKVVRVRPENTKEKAAGFKHVHVVDKHLVDFPKQEEVSFPFHGKTTNQNVIKQN-- 61

QY 60 ETLKYQDFPYGERSTQODIYAGSVOPILRHLEGGONASVLAYGPTGAGKTHMLGSPEQ 119
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 KDLKFVDFADVDETSTQSEVFEHTTAPILRSPLNGYNCVTLAYGATGAGKTHMLGSADE 121

QY 120 PGVPRALMDLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDPASGDLVIREDCRG 179
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 PGVWYLTMLHLKCMDEIKEE---ICSTAVSYLEVYNEQIRDLL-VNSGPLAVREDTQK 177

QY 180 NILIPGLS-QKPISSPADFERHFLP-ASRNRVTGATRLNQRSSRSRSHAVLLVKVDQERLA 237
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 178 GVVVHGLTLHQPSS--EETLHLLDNGKNRTOHPTDMNATSSRSRSHAVFQIYLRQODKTA 235

QY 238 PFRQ--REGKLYLIDLAGEEDNRRTGNKGLRKESCAINTSLFVLGKVVDAL----- 287
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 236 SINQNVRIAKMSLIDLAGESTSGAKGTRFVEGTNINRSLLALGNVINALADSKISFL 295

QY 288 -----NQGLPRVPYRDSKLTLLQDSLGGSAAHSILIANIAPERFYLDTVSALNFAARSK 342
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 296 CFQRKNQ---HIPYRNSKLTLLKDSLGNCQTIMIAAVSPSSVFYDDTYNTLKYANRAK 352

QY 343 EV 344
||: : : : :
Db 353 DI 354

RESULT 13
US-09-724-511-2
; Sequence 2, Application US/09724511
; Patent No. 6391601
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6391601el motor proteins and methods for
; FILE REFERENCE: 1046
; CURRENT APPLICATION NUMBER: US/09/724,511
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/632,344
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
```

```
Db 353 DI 354

RESULT 12
US-09-632-344-4
; Sequence 4, Application US/09632344
; Patent No. 6534309
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6534309el motor proteins and methods for
; FILE REFERENCE: 1046
; CURRENT APPLICATION NUMBER: US/09/632,344
; CURRENT FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Human
US-09-632-344-4

Query Match 31.1%; Score 549; DB 4; Length 355;
Best Local Similarity 39.0%; Pred. No. 4.6e-50;
Matches 141; Conservative 59; Mismatches 114; Indels 48; Gaps 11;

QY 20 VRVAVRLRP-----FVDGTAGASDP-----PCVRGMDSCSLEIANWRNHQ 59
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 MKVVRVRPENTKEKAAGFKHVHVVDKHLVDFPKQEEVSFPFHGKTTNQNVIKQN-- 61

QY 60 ETLKYQDFPYGERSTQODIYAGSVOPILRHLEGGONASVLAYGPTGAGKTHMLGSPEQ 119
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 KDLKFVDFADVDETSTQSEVFEHTTAPILRSPLNGYNCVTLAYGATGAGKTHMLGSADE 121

QY 120 PGVPRALMDLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDPASGDLVIREDCRG 179
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 PGVWYLTMLHLKCMDEIKEE---ICSTAVSYLEVYNEQIRDLL-VNSGPLAVREDTQK 177

QY 180 NILIPGLS-QKPISSPADFERHFLP-ASRNRVTGATRLNQRSSRSRSHAVLLVKVDQERLA 237
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 178 GVVVHGLTLHQPSS--EETLHLLDNGKNRTOHPTDMNATSSRSRSHAVFQIYLRQODKTA 235

QY 238 PFRQ--REGKLYLIDLAGEEDNRRTGNKGLRKESCAINTSLFVLGKVVDAL----- 287
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 236 SINQNVRIAKMSLIDLAGESTSGAKGTRFVEGTNINRSLLALGNVINALADSKISFL 295

QY 288 -----NQGLPRVPYRDSKLTLLQDSLGGSAAHSILIANIAPERFYLDTVSALNFAARSK 342
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 296 CFQRKNQ---HIPYRNSKLTLLKDSLGNCQTIMIAAVSPSSVFYDDTYNTLKYANRAK 352

QY 343 EV 344
||: : : : :
Db 353 DI 354

RESULT 13
US-09-724-511-2
; Sequence 2, Application US/09724511
; Patent No. 6391601
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6391601el motor proteins and methods for
; FILE REFERENCE: 1046
; CURRENT APPLICATION NUMBER: US/09/724,511
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/632,344
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
```

```
; SEQ ID NO 2
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Human
US-09-724-511-2

Query Match
 31.1%; Score 549; DB 3; Length 367;
Best Local Similarity 39.0%; Pred. No. 4.9e-50;
Matches 141; Conservative 59; Mismatches 114; Indels 48; Gaps 11;

QY 20 VRVAVRLRP-----FVDTAGASDP-----PCVRGMDSCSLEIANWRNHQ 59
Db 12 MKVVVRVPENTKEKAAGFKVHVVDKHLVDFDPKQEEVSFFHGKKTNNQVIKKQ-- 69
QY 60 ETLKYQDAFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAYGPTGAGKTHMLGSPEQ 119
Db 70 KDLKFVDAFVDETSTQSEVFEHTTKPIRLSFLNGYNTVLAYGATGAGKTHMLGSADE 129
QY 120 PGVIPRALMDLLQLTREBGAERPWALSVTMSYLEIYOEKVLDDLDPASGDLVIREDCRG 179
Db 130 PGVMYLTMLHLYKCMDEIKEE---ICSTAVSYLEVYNEQIRDLL--VNSGFLAVREDTQK 185
QY 180 NILIPGLS--QKPISSFADFERHFLP--ASRNTVGTATRLNQSRSSSHAVLLVKVDQERLA 237
Db 186 GVVVHGLTLHQPSS--EEILHLLDNGNKNRTOHPTDMNATSSRSHAVFQIYLRQODKTA 243
QY 238 PFRQ--REGKLYLIDLAGSEDNRRTGNKGLRLKESGAINTSIFVLGKVVVDAL----- 287
Db 244 SINQNVRIAKMSLIDLAGSERASTSGAKGTRFVEGTNINRSLALGNVINALADSKISFL 303
QY 288 -----NOGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSK 342
Db 304 CFQRKNQ---HIPYRNSKLTLLKDSLGNCQOTIMAAVSPSSVFYDDTYNTLYANRAK 360
QY 343 EV 344
Db 361 DI 362

RESULT 15
US-09-632-344-2
; Sequence 2, Application US/09632344
; Patent No. 6534309
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6534309el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1046
; CURRENT APPLICATION NUMBER: US/09/632,344
; CURRENT FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Human
US-09-632-344-2

Query Match
 31.1%; Score 549; DB 4; Length 367;
Best Local Similarity 39.0%; Pred. No. 4.9e-50;
Matches 141; Conservative 59; Mismatches 114; Indels 48; Gaps 11;

QY 20 VRVAVRLRP-----FVDTAGASDP-----PCVRGMDSCSLEIANWRNHQ 59
Db 12 MKVVVRVPENTKEKAAGFKVHVVDKHLVDFDPKQEEVSFFHGKKTNNQVIKKQ-- 69
QY 60 ETLKYQDAFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAYGPTGAGKTHMLGSPEQ 119
Db 70 KDLKFVDAFVDETSTQSEVFEHTTKPIRLSFLNGYNTVLAYGATGAGKTHMLGSADE 129
QY 120 PGVIPRALMDLLQLTREBGAERPWALSVTMSYLEIYOEKVLDDLDPASGDLVIREDCRG 179
Db 130 PGVMYLTMLHLYKCMDEIKEE---ICSTAVSYLEVYNEQIRDLL--VNSGFLAVREDTQK 185
QY 180 NILIPGLS--QKPISSFADFERHFLP--ASRNTVGTATRLNQSRSSSHAVLLVKVDQERLA 237
Db 186 GVVVHGLTLHQPSS--EEILHLLDNGNKNRTOHPTDMNATSSRSHAVFQIYLRQODKTA 243
QY 238 PFRQ--REGKLYLIDLAGSEDNRRTGNKGLRLKESGAINTSIFVLGKVVVDAL----- 287
Db 244 SINQNVRIAKMSLIDLAGSERASTSGAKGTRFVEGTNINRSLALGNVINALADSKISFL 303
QY 288 -----NOGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSK 342
Db 304 CFQRKNQ---HIPYRNSKLTLLKDSLGNCQOTIMAAVSPSSVFYDDTYNTLYANRAK 360
QY 343 EV 344
Db 361 DI 362

RESULT 14
US-09-723-097-2
; Sequence 2, Application US/09723097
; Patent No. 6492151
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6492151el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1046
; CURRENT APPLICATION NUMBER: US/09/723,097
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/632,344
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Human
US-09-723-097-2

Query Match
 31.1%; Score 549; DB 4; Length 367;
Best Local Similarity 39.0%; Pred. No. 4.9e-50;
Matches 141; Conservative 59; Mismatches 114; Indels 48; Gaps 11;

QY 20 VRVAVRLRP-----FVDTAGASDP-----PCVRGMDSCSLEIANWRNHQ 59
Db 12 MKVVVRVPENTKEKAAGFKVHVVDKHLVDFDPKQEEVSFFHGKKTNNQVIKKQ-- 69
QY 60 ETLKYQDAFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAYGPTGAGKTHMLGSPEQ 119
Db 70 KDLKFVDAFVDETSTQSEVFEHTTKPIRLSFLNGYNTVLAYGATGAGKTHMLGSADE 129
```

Search completed: November 5, 2004, 18:46:46
Job time : 14.5172 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 ; Search time 19.0257 Seconds
(without alignments)
1697.543 Million cell updates/sec

Title: US-10-797-893-8

Perfect score: 2472

Sequence: 1 MGRCLSKIGATRRPPPPARV.....LEAKMLAKAEKENHCPTM 487

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCFUS COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2472	100.0	487	3	US-09-724-224-8
2	2472	100.0	487	4	US-10-093-317-8
3	2467	99.8	512	3	US-09-724-224-4
4	2467	99.8	512	4	US-10-093-317-4
5	2467	99.8	665	4	US-09-595-684B-35
6	1768	71.5	346	3	US-09-724-224-6
7	1768	71.5	346	4	US-10-093-317-6
8	1763	71.3	370	3	US-09-724-224-2
9	1763	71.3	370	4	US-10-093-317-2
10	573.5	23.2	1234	4	US-09-592-054-8
11	572	23.1	522	4	US-09-592-054-4
12	568.5	23.0	473	4	US-09-592-054-6
13	561.5	22.7	1232	4	US-09-592-054-2
14	561	22.7	1388	3	US-09-572-191-2
15	561	22.7	1388	3	US-09-723-262-2
16	556.5	22.7	1388	3	US-09-723-219-2
17	556.5	22.5	1066	3	US-09-541-782-8
18	556.5	22.5	1066	4	US-09-723-820-8
19	556.5	22.5	1066	4	US-10-270-085-8
20	554	22.4	1279	3	US-09-724-517-2
21	554	22.4	1279	4	US-09-641-807A-2
22	554	22.4	1279	4	US-09-723-096-2
23	553.5	22.4	1231	4	US-09-595-684B-23
24	551	22.3	2954	4	US-09-150-867-1
25	550	22.2	864	4	US-09-883-096-2
26	549	22.2	355	3	US-09-724-511-4
27	549	22.2	355	4	US-09-723-097-4

28	549	22.2	355	4	US-09-632-344-4	Sequence 4, Appli
29	549	22.2	367	3	US-09-724-511-2	Sequence 2, Appli
30	549	22.2	367	4	US-09-723-097-2	Sequence 2, Appli
31	549	22.2	367	4	US-09-632-344-2	Sequence 2, Appli
32	546.5	22.1	1103	3	US-09-162-373-1	Sequence 1, Appli
33	546.5	22.1	1103	3	US-09-467-946-1	Sequence 25, Appli
34	544.5	22.0	935	4	US-09-914-259-25	Sequence 6, Appli
35	543	22.0	409	3	US-09-572-191-6	Sequence 6, Appli
36	543	22.0	409	3	US-09-723-262-6	Sequence 6, Appli
37	543	22.0	409	3	US-09-723-219-6	Sequence 6, Appli
38	543	22.0	928	4	US-09-914-259-23	Sequence 23, Appli
39	541	21.9	1690	4	US-09-595-684B-39	Sequence 39, Appli
40	535	21.6	513	4	US-09-724-519-6	Sequence 6, Appli
41	535	21.6	513	4	US-09-592-037-6	Sequence 6, Appli
42	535	21.6	513	4	US-09-428-156B-6	Sequence 2, Appli
43	533.5	21.5	1057	4	US-09-428-156B-2	Sequence 29, Appli
44	532	21.5	1056	4	US-09-595-684B-29	Sequence 10, Appli
45	532	21.5	1057	3	US-09-541-782-10	

ALIGNMENTS

RESULT 1
US-09-724-224-8
; Sequence 8, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09/724,224
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Human
US-09-724-224-8

Query Match	100.0%	Score	2472;	DB	3;	Length	487;
Best Local Similarity	100.0%	Pred. No.	1.1e-225;				
Matches	487;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MGRCLSKIGATRRPPPPARVAVRLRPFDVGTAGADPPCVRGMDSCSLEIANWRHNE	60				
Db	1	MGRCLSKIGATRRPPPPARVAVRLRPFDVGTAGADPPCVRGMDSCSLEIANWRHNE	60				
Qy	61	TLKYQDAFYGERSTQDDIYAGSVQPIRLHLLGQNASVLAAYGTGAGKTHMLGSPEQP	120				
Db	61	TLKYQDAFYGERSTQDDIYAGSVQPIRLHLLGQNASVLAAYGTGAGKTHMLGSPEQP	120				
Qy	121	GVTPRALMDLLQITRECAEGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN	180				
Db	121	GVTPRALMDLLQITRECAEGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN	180				
Qy	181	ILIPGLSOKPISGFADPFERHFLPASNRRTVGATRLNORSSRSHAVLVKVDQERLAPFR	240				
Db	181	ILIPGLSOKPISGFADPFERHFLPASNRRTVGATRLNORSSRSHAVLVKVDQERLAPFR	240				
Qy	241	QREGKLYLIDLAGSEDNRRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVPRYRDSK	300				
Db	241	QREGKLYLIDLAGSEDNRRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVPRYRDSK	300				
Qy	301	LTRLQDSLGGSAAHSILIANIAPERRFYLDVTSALNFAARKEVINRPFNTNESLQPHALG	360				
Db	301	LTRLQDSLGGSAAHSILIANIAPERRFYLDVTSALNFAARKEVINRPFNTNESLQPHALG	360				
Qy	361	PVKLSQKELLGPEAKRARGPEEEIGSPFMAAPASASQKLSPLQKLSMDPAMLERLL	420				

Db 361 PVKLSQKELLGPEAKRGPPEEEIGSPPEMAAPASQKLSPLQKLSMDPAMLERLL 420
 QY 421 SLDRLLASQSGCAPLLSTPKRRMVLMTVEKDLIEIRLTKKQLEAKMLAQKAEK 480
 Db 421 SLDRLLASQSGCAPLLSTPKRRMVLMTVEKDLIEIRLTKKQLEAKMLAQKAEK 480
 QY 481 ENHCPTM 487
 Db 481 ENHCPTM 487

RESULT 2
 US-10-093-317-8
 ; Sequence 8, Application US/10093317
 ; Patent No. 6762043
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; TITLE OF INVENTION: motor proteins and methods for
 ; TITLE OF INVENTION: their use
 ; FILE REFERENCE: 1044
 ; CURRENT APPLICATION NUMBER: US/10/093,317
 ; CURRENT FILING DATE: 2002-03-06
 ; PRIOR APPLICATION NUMBER: 09/724,224
 ; PRIOR FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 487
 ; TYPE: PRT
 ; ORGANISM: Human
 US-10-093-317-8

Query Match 100.0%; Score 2472; DB 4; Length 487;
 Best Local Similarity 100.0%; Pred. No. 1.1e-225;
 Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGRCLSKIGATRRPPARVAVRLRPFDVGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
 Db 1 MGRCLSKIGATRRPPARVAVRLRPFDVGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
 QY 61 TLKYQDAFYGERSTQODIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHMTLGSPEQ 120
 Db 61 TLKYQDAFYGERSTQODIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHMTLGSPEQ 120
 QY 121 GVIPRALMDLLQLTREBEGAEGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRG 180
 Db 121 GVIPRALMDLLQLTREBEGAEGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRG 180
 QY 181 ILIPGLSOKPISSFADPERHFLPASNRRTVGATRLNQRSRSHAVLLVKVDQERLAPR 240
 Db 181 ILIPGLSOKPISSFADPERHFLPASNRRTVGATRLNQRSRSHAVLLVKVDQERLAPR 240
 QY 241 QREGKLYLDLAGSEDNRRGTGKGLKESGAINTSFLVLGKVVDALNOGLPRVPYRDSK 300
 Db 241 QREGKLYLDLAGSEDNRRGTGKGLKESGAINTSFLVLGKVVDALNOGLPRVPYRDSK 300
 QY 301 LTRLLQDSLGSSAHSILIANIAPERRFYLDTSALNFAARSKVINRPTNESLQPHALG 360
 Db 301 LTRLLQDSLGSSAHSILIANIAPERRFYLDTSALNFAARSKVINRPTNESLQPHALG 360
 QY 361 PVKLSQKELLGPEAKRGPPEEEIGSPPEMAAPASQKLSPLQKLSMDPAMLERLL 420
 Db 361 PVKLSQKELLGPEAKRGPPEEEIGSPPEMAAPASQKLSPLQKLSMDPAMLERLL 420
 QY 421 SLDRLLASQSGCAPLLSTPKRRMVLMTVEKDLIEIRLTKKQLEAKMLAQKAEK 480
 Db 421 SLDRLLASQSGCAPLLSTPKRRMVLMTVEKDLIEIRLTKKQLEAKMLAQKAEK 480
 QY 481 ENHCPTM 487
 Db 481 ENHCPTM 487

RESULT 3
 US-09-724-224-4
 ; Sequence 4, Application US/09724224
 ; Patent No. 6387644
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; TITLE OF INVENTION: motor proteins and methods for
 ; TITLE OF INVENTION: their use
 ; FILE REFERENCE: 1044
 ; CURRENT APPLICATION NUMBER: US/09/724,224
 ; CURRENT FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: 09/597,292
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 512
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-724-224-4

Query Match 99.8%; Score 2467; DB 3; Length 512;
 Best Local Similarity 100.0%; Pred. No. 3.7e-225;
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GRCRLSKIGATRRPPARVAVRLRPFDVGTAGASDPPCVRGMDSCSLEIANWRNHQ 61
 Db 27 GRCRLSKIGATRRPPARVAVRLRPFDVGTAGASDPPCVRGMDSCSLEIANWRNHQ 86
 QY 62 LKYQDAFYGERSTQODIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHMTLGSPEQ 121
 Db 87 LKYQDAFYGERSTQODIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHMTLGSPEQ 146
 QY 122 VIPRALMDLLQLTREBEGAEGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRG 181
 Db 147 VIPRALMDLLQLTREBEGAEGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRG 206
 QY 182 LIPGLSOKPISSFADPERHFLPASNRRTVGATRLNQRSRSHAVLLVKVDQERLAPR 241
 Db 207 LIPGLSOKPISSFADPERHFLPASNRRTVGATRLNQRSRSHAVLLVKVDQERLAPR 266
 QY 242 REGKLYLDLAGSEDNRRGTGKGLKESGAINTSFLVLGKVVDALNOGLPRVPYRDSK 301
 Db 267 REGKLYLDLAGSEDNRRGTGKGLKESGAINTSFLVLGKVVDALNOGLPRVPYRDSK 326
 QY 302 LTRLLQDSLGSSAHSILIANIAPERRFYLDTSALNFAARSKVINRPTNESLQPHALG 361
 Db 327 LTRLLQDSLGSSAHSILIANIAPERRFYLDTSALNFAARSKVINRPTNESLQPHALG 386
 QY 362 VKLSQKELLGPEAKRGPPEEEIGSPPEMAAPASQKLSPLQKLSMDPAMLERLLS 421
 Db 387 VKLSQKELLGPEAKRGPPEEEIGSPPEMAAPASQKLSPLQKLSMDPAMLERLLS 446
 QY 422 LDRLLASQSGCAPLLSTPKRRMVLMTVEKDLIEIRLTKKQLEAKMLAQKAEK 481
 Db 447 LDRLLASQSGCAPLLSTPKRRMVLMTVEKDLIEIRLTKKQLEAKMLAQKAEK 506
 QY 482 NHCPTM 487
 Db 507 NHCPTM 512

RESULT 4
 US-10-093-317-4
 ; Sequence 4, Application US/10093317
 ; Patent No. 6762043
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; TITLE OF INVENTION: motor proteins and methods for
 ; TITLE OF INVENTION: their use
 ; FILE REFERENCE: 1044
 ; CURRENT APPLICATION NUMBER: US/10/093,317

; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Human
US-10-093-317-4

Query Match 99.8%; Score 2467; DB 4; Length 512;
Best Local Similarity 100.0%; Pred. No. 3.7e-225;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVAVLRPFVDTAGASDPPCVRGMDSCSLEIANRNHOET 61
DB 27 GRCLSKIGATRRPPPARVAVLRPFVDTAGASDPPCVRGMDSCSLEIANRNHOET 86

QY 62 LKQFDIFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLAAYGPTGAGKTHMTLGSPEQPG 121
DB 87 LKQFDIFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLAAYGPTGAGKTHMTLGSPEQPG 146

QY 122 VIPALMDLLQLTREAGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 181
DB 147 VIPALMDLLQLTREAGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 206

QY 182 LIPCLSKQPISSPADFERHFLPASRNTVGATRLNQRSSRSHAVLLVKVDORERLAPRQ 241
DB 207 LIPCLSKQPISSPADFERHFLPASRNTVGATRLNQRSSRSHAVLLVKVDORERLAPRQ 266

QY 242 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVVDALNQGLPRVPYRDSKL 301
DB 267 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVVDALNQGLPRVPYRDSKL 326

QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPTNESLOPHALGP 361
DB 327 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPTNESLOPHALGP 386

QY 362 VKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 421
DB 387 VKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 446

QY 422 LDRLLASQSGGAPLLSTPKRERVMKMTVEEKDLEIRLTKQKELEAKMLAQKABEKE 481
DB 447 LDRLLASQSGGAPLLSTPKRERVMKMTVEEKDLEIRLTKQKELEAKMLAQKABEKE 506

QY 482 NHCPTM 487
DB 507 NHCPTM 512

RESULT 5
US-09-595-684B-35
; Sequence 35, Application US/09595684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Ohashi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; FILE REFERENCE: cyto036
; CURRENT APPLICATION NUMBER: US/09/595,684B
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35

; LENGTH: 665
; TYPE: PRT
; ORGANISM: Human
US-09-595-684B-35

Query Match 99.8%; Score 2467; DB 4; Length 665;
Best Local Similarity 100.0%; Pred. No. 5.6e-225;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVAVLRPFVDTAGASDPPCVRGMDSCSLEIANRNHOET 61
DB 26 GRCLSKIGATRRPPPARVAVLRPFVDTAGASDPPCVRGMDSCSLEIANRNHOET 85

QY 62 LKQFDIFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLAAYGPTGAGKTHMTLGSPEQPG 121
DB 86 LKQFDIFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLAAYGPTGAGKTHMTLGSPEQPG 145

QY 122 VIPALMDLLQLTREAGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 181
DB 146 VIPALMDLLQLTREAGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 205

QY 182 LIPCLSKQPISSPADFERHFLPASRNTVGATRLNQRSSRSHAVLLVKVDORERLAPRQ 241
DB 206 LIPCLSKQPISSPADFERHFLPASRNTVGATRLNQRSSRSHAVLLVKVDORERLAPRQ 265

QY 242 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVVDALNQGLPRVPYRDSKL 301
DB 266 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVVDALNQGLPRVPYRDSKL 325

QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPTNESLOPHALGP 361
DB 326 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPTNESLOPHALGP 385

QY 362 VKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 421
DB 386 VKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 445

QY 422 LDRLLASQSGGAPLLSTPKRERVMKMTVEEKDLEIRLTKQKELEAKMLAQKABEKE 481
DB 446 LDRLLASQSGGAPLLSTPKRERVMKMTVEEKDLEIRLTKQKELEAKMLAQKABEKE 505

QY 482 NHCPTM 487
DB 506 NHCPTM 511

RESULT 6
US-09-724-224-6
; Sequence 6, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
; FILE REFERENCE: their use
; CURRENT APPLICATION NUMBER: US/09/724,224
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Human
US-09-724-224-6

Query Match 71.5%; Score 1768; DB 3; Length 346;
Best Local Similarity 100.0%; Pred. No. 4e-159;
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRCLSKIGATRRPPPARVAVLRPFVDTAGASDPPCVRGMDSCSLEIANRNHOE 60

Db 1 MGRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHOE 60
QY 61 TLKYQFDAYFGRSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHTMLGSPROP 120
Db 61 TLKYQFDAYFGRSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHTMLGSPROP 120
QY 121 GVTPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180
Db 121 GVTPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180
QY 181 ILIPGLSQPISSFADFERRHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR 240
Db 181 ILIPGLSQPISSFADFERRHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR 240
QY 241 QREGKLYLIDLAGESEDNRRRTGNKGLRKESGAINTSLFVLGKVVADALNOGLPRVPYRDSK 300
Db 241 QREGKLYLIDLAGESEDNRRRTGNKGLRKESGAINTSLFVLGKVVADALNOGLPRVPYRDSK 300
QY 301 LTRLLQDSLGGSASHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346
Db 301 LTRLLQDSLGGSASHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346

RESULT 7

US-10-093-317-6
; Sequence 6, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093,317
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Human
US-10-093-317-6

Query Match 71.5%; Score 1768; DB 4; Length 346;
Best Local Similarity 100.0%; Pred. No. 4e-159;
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHOE 60
Db 1 MGRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHOE 60
QY 61 TLKYQFDAYFGRSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHTMLGSPROP 120
Db 61 TLKYQFDAYFGRSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHTMLGSPROP 120
QY 121 GVTPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180
Db 121 GVTPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180
QY 181 ILIPGLSQPISSFADFERRHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR 240
Db 181 ILIPGLSQPISSFADFERRHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR 240
QY 241 QREGKLYLIDLAGESEDNRRRTGNKGLRKESGAINTSLFVLGKVVADALNOGLPRVPYRDSK 300
Db 241 QREGKLYLIDLAGESEDNRRRTGNKGLRKESGAINTSLFVLGKVVADALNOGLPRVPYRDSK 300
QY 301 LTRLLQDSLGGSASHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346
Db 301 LTRLLQDSLGGSASHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346

RESULT 8

US-09-724-224-2
; Sequence 2, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09/724,224
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Human
US-09-724-224-2

Query Match 71.3%; Score 1763; DB 3; Length 370;
Best Local Similarity 100.0%; Pred. No. 1.3e-158;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCRLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHOE 61
Db 26 GRCRLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHOE 85
QY 62 LKYQFDAYFGRSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHTMLGSPROP 121
Db 86 LKYQFDAYFGRSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHTMLGSPROP 145
QY 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 181
Db 146 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 205
QY 182 LIPGLSQPISSFADFERRHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR 241
Db 206 LIPGLSQPISSFADFERRHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR 265
QY 242 REGKLYLIDLAGESEDNRRRTGNKGLRKESGAINTSLFVLGKVVADALNOGLPRVPYRDSK 301
Db 266 REGKLYLIDLAGESEDNRRRTGNKGLRKESGAINTSLFVLGKVVADALNOGLPRVPYRDSK 325
QY 302 TRLLODSLGGSASHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346
Db 326 TRLLODSLGGSASHSILIANIAPERRFYLDTVSALNFAARSKEVIN 370

RESULT 9

US-10-093-317-2
; Sequence 2, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093,317
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Human
US-10-093-317-2

Query Match 71.3%; Score 1763; DB 4; Length 370;
Best Local Similarity 100.0%; Pred. No. 1.3e-158;

Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRCLSKIGATRRPPRRVAVRLRPVDTAGASDPPCVRGMDSCSLEIANWRNHQET 61
Db
26 GRCLSKIGATRRPPRRVAVRLRPVDTAGASDPPCVRGMDSCSLEIANWRNHQET 85
Qy 62 LKYPDAPYGRSTQODIYAGSVQPIHLHLLLEGONASVLYAGPTGAGKTHMLGSPQPG 121
Db 86 LKYPDAPYGRSTQODIYAGSVQPIHLHLLLEGONASVLYAGPTGAGKTHMLGSPQPG 145
Qy 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDVIREDCRNI 181
Db 146 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDVIREDCRNI 205
Qy 182 LIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 241
Db 206 LIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 265
Qy 242 REGKYLIDLAGSEDRNRGTGKGLRKESGAINSTSLFVLGKVDALNQGRLPRVPRYRDSKL 301
Db 266 REGKYLIDLAGSEDRNRGTGKGLRKESGAINSTSLFVLGKVDALNQGRLPRVPRYRDSKL 325
Qy 302 TRLLQDSIGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346
Db 326 TRLLQDSIGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 370

RESULT 10

US-09-592-054-8

; Sequence 8, Application US/09592054

; Patent No. 6440684

; GENERAL INFORMATION:

; APPLICANT: Beraud, Christophe

; APPLICANT: Finer, Jeffrey

; APPLICANT: Sakowicz, Roman

; APPLICANT: Wood, Kenneth

; TITLE OF INVENTION: No. 6440684el motor proteins and methods for

; FILE REFERENCE: 1016

; CURRENT APPLICATION NUMBER: US/09/592,054

; CURRENT FILING DATE: 2000-07-20

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 1234

; TYPE: PRT

; ORGANISM: Human

US-09-592-054-8

Query Match 23.2%; Score 573.5; DB 4; Length 1234;
Best Local Similarity 32.2%; Pred. No. 6.1e-45;
Matches 156; Conservative 97; Mismatches 173; Indels 59; Gaps 13;

Qy 20 VRVAVRLRPVDTAGASDPPC-----VRGMDSCSLEIANWRNHQETLKYQDAPF 69
Db 10 VRVALRCRPLVPKEISEGCQCLSFVPGETQVVVGTDK-----SFTYDFV 54
Qy 70 YGERSTQODIYAGSVQPIHLHLLLEGONASVLYAGPTGAGKTHMLGS-----PEQP--GV 122
Db 55 FDPCTEQEYFNVKAVAPLIGIFKGYNATVLYAGTGTSGKTSYMGGAAYTAQEONEPTVGI 114
Qy 123 IPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPA--SGDLVIREDCRNI 180
Db 115 IPR-----VIQLLKEIDQKDFEFTLVKSVLEIYNEIILDLCPREKAQINIREDPKEG 170
Qy 181 ILIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR 240
Db 171 IKIVGLTEKTVLVALDVTSCLEQGNRSRTVASTAMNSQSSRSHAIPTISLEQCKGSKNS 230
Qy 241 QREGKYLIDLAGSEDRNRGTGKGLRKESGAINSTSLFVLGKVDALNQGRLPR--VPYRD 298
Db 231 SFRSKHLVLAGSERQKTKAEGDRLKEGININRGLLCLGNVISALGDDKKGKGFVPRD 290

Qy 299 SKLTRLQDSIGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPTNBSLOPHA 358
Db 291 SKLTRLQDSIGSSAHSILIANIACVSPADSNLEETLSTLYADARKKIKNKPVN--IDPHT 348
Qy 359 LGPVLSQKELLGPEAKRARGPEEEIEGSPPEPMAAPASASQKLSPLQKLSMDPAMLER 418
Db 349 AELNHLKQ-----QVQLQVLLQALHGGTLPGSINAEPSEN---LQSLMEKNGSLVEE 398
Qy 419 LLSLDRLLASQGSQAGPLSTPKRERMVLMKTVEEK-DLETERLKTQK-ELEAKMLAQK 476
Db 399 NEKLSRCLSKAAGTAQML-----ERILTEQVNEKLNKLEELRQHAACKLDLQKUVET 453
Qy 477 AEEKE 481
Db 454 LEDQE 458

RESULT 11

US-09-592-054-4

; Sequence 4, Application US/09592054

; Patent No. 6440684

; GENERAL INFORMATION:

; APPLICANT: Beraud, Christophe

; APPLICANT: Finer, Jeffrey

; APPLICANT: Sakowicz, Roman

; APPLICANT: Wood, Kenneth

; TITLE OF INVENTION: No. 6440684el motor proteins and methods for

; FILE REFERENCE: 1016

; CURRENT APPLICATION NUMBER: US/09/592,054

; CURRENT FILING DATE: 2000-07-20

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 522

; TYPE: PRT

; ORGANISM: Human

US-09-592-054-4

Query Match 23.1%; Score 572; DB 4; Length 522;
Best Local Similarity 31.8%; Pred. No. 2.1e-45;
Matches 165; Conservative 99; Mismatches 167; Indels 88; Gaps 17;

Qy 1 MGRCLSKIGATRRPPA-----RVRAVRLRP-----FVDTAGASD 38
Db 10 MGRIRI-----RAPSTSLVEEVGIPVRVALRCRPLVPKEISEGCQCLSFVPG-----E 58
Qy 39 PPCVRGMDCSLEIANWRNHQETLKYQDAPYGERSTQODIYAGSVQPIHLHLLLEGONAS 98
Db 59 PQVVVGTDK-----SFTYDFVDPSTEQEYFNTAVAPLIGVFKGYNAT 103
Qy 99 VLYAGPTGAGKTHMLGS-----PEQP--GVIPRALMDLLQLTREEGAGRPWALSVTMS 151
Db 104 VLYAGTGTSGKTSYMGGAAYTAQEONEPTVGI PR-----VIQLLKEIDKSKSDFEFTLKVS 159
Qy 152 YLSIYQEKVLDLDPA--SGDLVIREDCRNIILPGLSQKPISSFADFERHFLPASNRRT 209
Db 160 YLSIYNEIILDLCPREKAQINIREDPKEGILVGLTEKTVLVALDVTSCLEQGNRSRT 219
Qy 210 VGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGSEDRNRRTGKGLRKE 269
Db 220 VASTAMNSQSSRSHAIPTISLEQCKGSKNSFRSKLHLVDLAGSERQKTKAEGDRLKE 279
Qy 270 SGAINSTSLFVLGKVDALNQGRLPR--VPYRDSKLTRLQDSIGSSAHSILIANIAPERRP 327
Db 280 GININRGLLCLGNVISALGDDKKGKGFVPRYRDSKLTRLQDSIGSSHTLTIACVSPADSN 339
Qy 328 YLDTVSALNFAARSKEVINRPTNBSLOPHALGPVK--LSQKELGPEAKRARGPEEE 385
Db 340 LESTLTLNRYADARKKIKNKPVINIDPQTALNHLKQVQVQLV-----LLQAHGG----- 391
Qy 386 IGSPEPMAAPASASQKLSL--LQKLSNMDPAMLERLLSLDRLLASQGSQAGPLSTPKRER 444

```

392 -----TLPGSITVPESENLSQSLMEKQSLVEENEKLSRGLSEAAQTQAML-----ER 439
445 MVLMTVEEK-DLEIERLTKQK-ELEAKMLAQAEKE 481
440 IITWEQANEKNAKLEELRQHAACKLDLQKLVETLEDQE 478

RESULT 12
US-09-592-054-6
; Sequence 6, Application US/09592054
; Patent No. 6440684
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Finer, Jeffrey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6440684el motor proteins and methods for
; FILE REFERENCE: 1016
; CURRENT APPLICATION NUMBER: US/09/592,054
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Human
US-09-592-054-6

Query Match 23.0%; Score 568.5; DB 4; Length 473;
Best Local Similarity 32.3%; Pred. No. 3.9e-45;
Matches 159; Conservative 97; Mismatches 162; Indels 75; Gaps 15;

QY 20 VRVAVLRP-----FVDGTAGASDPPCVRGMDSCLSEIANWRNHQETLKY 64
DB 8 VRVALRCPLVPKEISGCMCLSFVPG-----EPQVVVGTDK-----SF 47
QY 65 QFDIFYGERSTQDDIYAGSVQPIRLHLEGONASVLAAGPTGAGKTHMLGS-----PEQ 119
DB 48 TYDFVDPSTQEVEVNTAVAPLKGKGYGNATVLAAYGQTSKTYSMGGAYTABQENE 107
QY 120 P--GVIPRALMDLLQTLREGEAGRPWALSVTMSYLEIYQEKVLDLDDPA--SGDLVIRE 175
DB 108 PTGVGIPR-----VIQLLFEIDKSDFEFTLKVSYLEIYNEEILDLCPREKAQINIRE 163
QY 176 DCRGNILPGLSQKPISSFADFERHFLPASNRVTGATRLNORSSRSHAVLLVKVDQER 235
DB 164 DPKEGKIVGLTEKTVLVALDVTSCLEQNNRSRTVASTAMNSQSSRSHAIPTISLEQKK 223
QY 236 LAPFRQREGKLYLDLAGSEDRNRGKGLRKESGAINSLFVLGKVDALNQLPR-- 293
DB 224 SDKNSFRSKLHLVDLAGSERQKTKAEGDRKKEGININRGLLCLGNVISALGDDKKGPF 283
QY 294 VPYRDSKLTLLQDSLGSGSAHSILIANIAPERRFYDVTVSALNFAARKEVINRPFTHES 353
DB 284 VPYRDSKLTLLQDSLGSGNSHTLMACVSPADSNLEETLNTLRYADRAKIKNKPINVID 343
QY 354 LQPHALGPVK--LSQKELLGPPEAKRGPPEEIGSPPEMAAPASASQKLSL-LQKLSS 410
DB 344 PQTAEHLNHLKQVQQLQVL-----LIQAHHG-----TLPGSITVPESENLSQSLME 389
QY 411 MDPAMLERLLSLDRLLASQSGCAPLLSTPKRERVMVLMKTVEEK-DLEIERLTKQK-EL 468
DB 389 KQSLVEENEKLSRGLSEAAQTQAML-----ERLIWTEQANEKNAKLEELRQHAACKL 443
QY 469 EAKMLAQAEKE 481
DB 444 DLQKLVETLEDQE 456

RESULT 13
US-09-592-054-2
; Sequence 2, Application US/09592054

```

```

; Patent No. 6440684
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Finer, Jeffrey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6440684el motor proteins and methods for
; FILE REFERENCE: 1016
; CURRENT APPLICATION NUMBER: US/09/592,054
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1232
; TYPE: PRT
; ORGANISM: Human
US-09-592-054-2

Query Match 22.7%; Score 561.5; DB 4; Length 1232;
Best Local Similarity 32.0%; Pred. No. 8.3e-44;
Matches 158; Conservative 97; Mismatches 163; Indels 75; Gaps 15;

QY 20 VRVAVLRP-----FVDGTAGASDPPCVRGMDSCLSEIANWRNHQETLKY 64
DB 10 VRVALRCPLVPKEISGCMCLSFVPG-----EPQVVVGTDK-----SF 49
QY 65 QFDIFYGERSTQDDIYAGSVQPIRLHLEGONASVLAAGPTGAGKTHMLGS-----PEQ 119
DB 50 TYDFVDPSTQEVEVNTAVAPLKGKGYGNATVLAAYGQTSKTYSMGGAYTABQENE 109
QY 120 P--GVIPRALMDLLQTLREGEAGRPWALSVTMSYLEIYQEKVLDLDDPA--SGDLVIRE 175
DB 110 PTGVGIPR-----VIQLLFEIDKSDFEFTLKVSYLEIYNEEILDLCPREKAQINIRE 165
QY 176 DCRGNILPGLSQKPISSFADFERHFLPASNRVTGATRLNORSSRSHAVLLVKVDQER 235
DB 166 DPKEGKIVGLTEKTVLVALDVTSCLEQNNRSRTVASTAMNSQSSRSHAIPTISLEQKK 225
QY 236 LAPFRQREGKLYLDLAGSEDRNRGKGLRKESGAINSLFVLGKVDALNQLPR-- 293
DB 226 SDKNSFRSKLHLVDLAGSERQKTKAEGDRKKEGININRGLLCLGNVISALGDDKKGPF 285
QY 294 VPYRDSKLTLLQDSLGSGSAHSILIANIAPERRFYDVTVSALNFAARKEVINRPFTHES 353
DB 286 APYRDSKLTLLQDSLGSGNSHTLMACVSPADSNLEETLNTLRYADRAKIKNKPINVID 345
QY 354 LQPHALGPVK--LSQKELLGPPEAKRGPPEEIGSPPEMAAPASASQKLSL-LQKLSS 410
DB 346 PQTAEHLNHLKQVQQLQVL-----LIQAHHG-----TLPGSITVPESENLSQSLME 390
QY 411 MDPAMLERLLSLDRLLASQSGCAPLLSTPKRERVMVLMKTVEEK-DLEIERLTKQK-EL 468
DB 391 KQSLVEENEKLSRGLSEAAQTQAML-----ERLIWTEQANEKNAKLEELRQHAACKL 445
QY 469 EAKMLAQAEKE 481
DB 446 DLQKLVETLEDQE 458

RESULT 14
US-09-572-191-2
; Sequence 2, Application US/09572191
; Patent No. 6355466
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6355466el motor proteins and methods for
; FILE REFERENCE: 1017
; CURRENT APPLICATION NUMBER: US/09/572,191
; CURRENT FILING DATE: 2000-05-17

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:44:52 ; Search time 82.6338 Seconds
(without alignments)
2081.726 Million cell updates/sec

Title: US-10-797-893-8

Perfect score: 2472

Sequence: 1 MGRCLSKIGATRRPPPARV.....LEAKMLAQKAEKENHCHPTM 487

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 35325886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09D_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US10G_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US10H_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2472	100.0	487	16	US-10-797-893-8
2	2467	99.8	512	16	US-10-797-893-4
3	2309	93.4	490	15	US-10-334-143-80
4	2141	86.6	460	9	US-09-925-300-1228
5	1768	71.5	346	16	US-10-797-893-6
6	1763	71.3	370	16	US-10-797-893-2
7	608	24.6	589	17	US-10-425-115-320209
8	601	24.3	377	16	US-10-437-963-159031
9	578.5	23.4	1237	15	US-10-334-143-33
10	574.5	23.2	1232	14	US-10-116-712-670
11	574.5	23.2	1232	16	US-10-408-765A-2153
12	574.5	23.2	1235	15	US-10-334-143-8
13	571	23.1	776	16	US-10-437-963-103943

14	571	23.1	783	17	US-10-425-115-332919	Sequence 332919,
15	571	23.1	854	15	US-10-425-114-59708	Sequence 59708, A
16	564.5	22.8	548	15	US-10-108-260A-2692	Sequence 2692, Ap
17	563.5	22.8	1001	17	US-10-425-115-231231	Sequence 231231,
18	561	22.7	1388	14	US-10-146-473-82	Sequence 82, Appl
19	561	22.7	1388	15	US-10-173-999-32	Sequence 32, Appl
20	561	22.7	1388	15	US-10-332-089-2	Sequence 2, Appl
21	561	22.7	1388	15	US-10-188-832-164	Sequence 164, App
22	560	22.7	992	17	US-10-739-930-7903	Sequence 7903, Ap
23	559.5	22.6	1401	15	US-10-287-226-142	Sequence 142, App
24	559	22.6	1030	15	US-10-425-114-62748	Sequence 62748, A
25	557.5	22.6	672	16	US-10-408-765A-1664	Sequence 1664, Ap
26	557.5	22.6	1232	14	US-10-116-712-664	Sequence 664, App
27	557.5	22.6	1232	14	US-10-116-712-669	Sequence 669, App
28	557	22.5	1324	15	US-10-287-226-314	Sequence 314, App
29	554	22.4	757	15	US-10-220-120-366	Sequence 366, App
30	554	22.4	757	16	US-10-363-829-316	Sequence 316, App
31	550	22.2	864	9	US-09-883-096-2	Sequence 2, Appl
32	546.5	22.1	1103	9	US-09-847-874A-1	Sequence 1, Appl
33	546.5	22.1	1103	14	US-10-458-162-1	Sequence 1, Appl
34	546.5	22.1	1826	16	US-10-408-765A-1477	Sequence 1477, Ap
35	545.5	22.1	834	15	US-10-287-226-306	Sequence 306, App
36	544.5	22.0	935	14	US-10-080-608A-25	Sequence 25, Appl
37	544.5	22.0	935	14	US-10-370-685-114	Sequence 114, App
38	543	22.0	409	15	US-10-332-089-6	Sequence 6, Appl
39	543	22.0	928	14	US-10-080-608A-23	Sequence 23, Appl
40	543	22.0	928	14	US-10-370-685-112	Sequence 112, App
41	541.5	21.9	1382	16	US-10-437-963-176714	Sequence 176714,
42	535	21.6	513	17	US-10-601-036-6	Sequence 6, Appl
43	533.5	21.6	1057	17	US-10-601-036-2	Sequence 2, Appl
44	532	21.5	1056	14	US-10-282-174-472	Sequence 472, App
45	532	21.5	1056	14	US-10-282-174-474	Sequence 474, App

ALIGNMENTS

RESULT 1
US-10-797-893-8
; Sequence 8, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10797,893
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Human
; US-10-797-893-8

Query Match 100.0%; Score 2472; DB 16; Length 487;
Best Local Similarity 100.0%; Pred No. 9.3e-202;
Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGRCLSKIGATRRPPPARVVRVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRHQE 60
Db 1 MGRCLSKIGATRRPPPARVVRVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRHQE 60
Qy 61 TLKQDFAFYGERSTOODIYAGSVQPIRLHLLRGONASVLAYGPTGAGKTHMLGSPQ 120
Db 61 TLKQDFAFYGERSTOODIYAGSVQPIRLHLLRGONASVLAYGPTGAGKTHMLGSPQ 120
Qy 121 GVIPRALMDLLQTLREAGRGPRWALSVTWTSYLEIYQKVLDDLDPASGDLVIREDCRGN 180

Db 121 GVIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQYKVDLLOPASGDLVIREDCRGN 180
 QY 181 ILIPGLSQPISFADFHEHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 240
 Db 181 ILIPGLSQPISFADFHEHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 240
 QY 241 QREGKLYLIDLAGESEDRNRRTGNKGLRLESKGAINTSLFVLGKVDALNOGLPRVPYRDSK 300
 Db 241 QREGKLYLIDLAGESEDRNRRTGNKGLRLESKGAINTSLFVLGKVDALNOGLPRVPYRDSK 300
 QY 301 LTRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPTNESLOPHALG 360
 Db 301 LTRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPTNESLOPHALG 360
 QY 361 PVKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420
 Db 361 PVKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420
 QY 421 SLDRLLASQGSQCAPLLSTPKRERMVLMKTVEKOLEIERLTKYKKELEAKMLAQKAEK 480
 Db 421 SLDRLLASQGSQCAPLLSTPKRERMVLMKTVEKOLEIERLTKYKKELEAKMLAQKAEK 480
 QY 481 ENHCPTM 487
 Db 481 ENHCPTM 487

RESULT 2

US-10-797-893-4
 ; Sequence 4, Application US/10797893
 ; Publication No. US2004014237A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; TITLE OF INVENTION: Novel motor proteins and methods for
 ; TITLE OF INVENTION: their use
 ; FILE REFERENCE: 1044
 ; CURRENT APPLICATION NUMBER: US/10797,893
 ; CURRENT FILING DATE: 2004-03-09
 ; PRIOR APPLICATION NUMBER: US/09/724,224
 ; PRIOR FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: 09/597,292
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 512
 ; TYPE: PRT
 ; ORGANISM: Human
 US-10-797-893-4

Query Match 99.8%; Score 2467; DB 16; Length 512;
 Best Local Similarity 100.0%; Pred. No. 2.7e-201;
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GRCRLSKIGATRRPPPARVRVAVLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61
 Db 27 GRCRLSKIGATRRPPPARVRVAVLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 86
 QY 62 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQPG 121
 Db 87 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQPG 146
 QY 122 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQYKVDLLOPASGDLVIREDCRGN 181
 Db 147 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQYKVDLLOPASGDLVIREDCRGN 206
 QY 182 LIPGLSQPISFADFHEHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 241
 Db 207 LIPGLSQPISFADFHEHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 266
 QY 242 REGKLYLIDLAGESEDRNRRTGNKGLRLESKGAINTSLFVLGKVDALNOGLPRVPYRDSK 301

Db 267 REGKLYLIDLAGESEDRNRRTGNKGLRLESKGAINTSLFVLGKVDALNOGLPRVPYRDSK 326
 QY 302 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPTNESLOPHALGP 361
 Db 327 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPTNESLOPHALGP 386
 QY 362 VKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 421
 Db 387 VKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 446
 QY 422 LDRLLASQGSQCAPLLSTPKRERMVLMKTVEKOLEIERLTKYKKELEAKMLAQKAEK 481
 Db 447 LDRLLASQGSQCAPLLSTPKRERMVLMKTVEKOLEIERLTKYKKELEAKMLAQKAEK 506
 QY 482 NHCPTM 487
 Db 507 NHCPTM 512

RESULT 3

US-10-334-143-80
 ; Sequence 80, Application US/10334143
 ; Publication No. US20040009549A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH
 ; APPLICANT: SUDARSANAM, SUCHA
 ; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
 ; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD
 ; FILE REFERENCE: 038602/1543
 ; CURRENT APPLICATION NUMBER: US/10/334,143
 ; CURRENT FILING DATE: 2002-12-31
 ; PRIOR APPLICATION NUMBER: 60/343,169
 ; PRIOR FILING DATE: 2001-12-31
 ; NUMBER OF SEQ ID NOS: 207
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 80
 ; LENGTH: 490
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-334-143-80

Query Match 93.4%; Score 2309; DB 15; Length 490;
 Best Local Similarity 100.0%; Pred. No. 7.4e-188;
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCRLSKIGATRRPPPARVRVAVLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61
 Db 34 GRCRLSKIGATRRPPPARVRVAVLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 93
 QY 62 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQPG 121
 Db 94 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQPG 153
 QY 122 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQYKVDLLOPASGDLVIREDCRGN 181
 Db 154 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQYKVDLLOPASGDLVIREDCRGN 213
 QY 182 LIPGLSQPISFADFHEHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 241
 Db 214 LIPGLSQPISFADFHEHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 273
 QY 242 REGKLYLIDLAGESEDRNRRTGNKGLRLESKGAINTSLFVLGKVDALNOGLPRVPYRDSK 301
 Db 274 REGKLYLIDLAGESEDRNRRTGNKGLRLESKGAINTSLFVLGKVDALNOGLPRVPYRDSK 333
 QY 302 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPTNESLOPHALGP 361
 Db 334 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPTNESLOPHALGP 393
 QY 362 VKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 421
 Db 394 VKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 453

Qy 422 LDRLLASQSGAPILLSTPKRERMVLMKTVBEKOL 456
Db 454 LDRLLASQSGAPILLSTPKRERMVLMKTVBEKOL 488

RESULT 4

US-09-925-300-1228
; Sequence 1228, Application US/09925300
; Publication No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925.300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1228
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (75)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (147)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (435)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1228

Query Match 86.6%; Score 2141; DB 9; Length 460;
Best Local Similarity 95.9%; Pred. No. 1.4e-173;
Matches 422; Conservative 1; Mismatches 17; Indels 0; Gaps 0;
Qy 2 GRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHOET 61
Db 4 GRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHOET 63
Qy 62 LKYQDFAFGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPRQP 121
Db 64 LKYQDFAFGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPRQP 123
Qy 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 181
Db 124 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 183
Qy 182 LIPGLSQKPISSPADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFRQ 241
Db 184 LIPGLSQKPISSPADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFRQ 243
Qy 242 REGKLYLIDLAGEEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDSKL 301
Db 244 REGKLYLIDLAGEEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDSKL 303
Qy 302 TRLLQDSLGSASHILIANIAPERRFYLDTVSALNFAARKEVINRPTNESLOPHALGP 361
Db 304 TRLLQDSLGSASHILIANIAPERRFYLDTVSALNFAARKEVINRPTNESLOPHALGP 363
Qy 362 VKLSQKELLGPPPEAKRGPPEEIGSPPEMAAPASQKLSPLQKLSMDPAMLERLLS 421
Db 364 VKLSQKELLGPPPEAKRGPPEEIGSPPEMAAPASQKLSPLQKLSMDPAMLERLLQ 423
Qy 422 LDRLLASQSGAPILLSTPK 441
Db 424 LGPSACLPGEPXGFSVEYPK 443

RESULT 5

US-10-797-893-6
; Sequence 6, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/797.893
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Human
US-10-797-893-6

Query Match 71.5%; Score 1768; DB 16; Length 346;
Best Local Similarity 100.0%; Pred. No. 5.8e-142;
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHOE 60
Db 1 MGRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHOE 60
Qy 61 TLKYQDFAFGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPRQP 120
Db 61 TLKYQDFAFGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPRQP 120
Qy 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180
Db 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180
Qy 181 ILIPGLSQKPISSPADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 240
Db 181 ILIPGLSQKPISSPADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 240
Qy 241 ORGKLYLIDLAGEEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDSK 300
Db 241 ORGKLYLIDLAGEEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDSK 300
Qy 301 LTRLLQDSLGSASHILIANIAPERRFYLDTVSALNFAARKEVIN 346
Db 301 LTRLLQDSLGSASHILIANIAPERRFYLDTVSALNFAARKEVIN 346

RESULT 6

US-10-797-893-2
; Sequence 2, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/797.893
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 370

```

: TYPE: PRT
: ORGANISM: Human
: US-10-797-893-2

Query Match
Best Local Similarity 71.3%; Score 1763; DB 16; Length 370;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCRLSKIGATRRPPPARVAVRLRPVDTAGASDPPCVGRGMDSCSLSIANWRNHQET 61
D 26 GRCRLSKIGATRRPPPARVAVRLRPVDTAGASDPPCVGRGMDSCSLSIANWRNHQET 85
QY 62 LKYQDFAGFGERSTQODIYAGSVQPIRLHLEGGONASVLAEGPTGAGKTHMLGSPQPG 121
D 86 LKYQDFAGFGERSTQODIYAGSVQPIRLHLEGGONASVLAEGPTGAGKTHMLGSPQPG 145
QY 122 VIPRALMDLLQLTREGAAGRPWALSVMYSLEYIQEYKVLDDLDPASGDVIREDCRGN 181
D 146 VIPRALMDLLQLTREGAAGRPWALSVMYSLEYIQEYKVLDDLDPASGDVIREDCRGN 205
QY 182 LIPGLSQPISFADFPERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFRQ 241
D 206 LIPGLSQPISFADFPERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFRQ 265
QY 242 REGKLYLIDLAGSEDRRTGNKGLRKESGAINTSFLVLGKVVADALNQGLPRVYRDSKL 301
D 266 REGKLYLIDLAGSEDRRTGNKGLRKESGAINTSFLVLGKVVADALNQGLPRVYRDSKL 325
QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDTSALNFAARSKEVIN 346
D 326 TRLLQDSLGSAHSILIANIAPERRFYLDTSALNFAARSKEVIN 370

RESULT 7
US-10-425-115-320209
: Sequence 320209, Application US/10425115
: Publication No. US20040214272A1
: GENERAL INFORMATION:
: APPLICANT: La Rosa, Thomas J.
: APPLICANT: Kovalic, David K.
: APPLICANT: Zhou, Yihua
: APPLICANT: Cao, Yongwei
: APPLICANT: Li, Ping
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
: TITLE OF INVENTION: Plants
: FILE REFERENCE: 38-21(53222)B
: CURRENT APPLICATION NUMBER: US/10/425,115
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 369326
: SEQ ID NO 320209
: LENGTH: 589
: TYPE: PRT
: ORGANISM: Zea mays
: FEATURE:
: OTHER INFORMATION: Clone ID: MRT4577_55097C.1.pep
US-10-425-115-320209

Query Match
Best Local Similarity 24.6%; Score 608; DB 17; Length 589;
Matches 174; Conservative 94; Mismatches 163; Indels 128; Gaps 19;

QY 17 PAR-----VRVAVRLRPVDTAGASDPPCVGRGMDSC-----SLSIANWRNHQETL 62
D 7 PARSAHLSQPVRLVRLVRLPFLSEASATPCV-SILGCHPGGGVTVLKD-QHTSRSE 64
QY 63 KYQDFAGFGERSTQODIYAGSVQPIRLHLEGGONASVLAEGPTGAGKTHMLGSPQPG 122
D 65 QYKLDAPFRQSDVSQIFDQEVRAVIPSIFEGINATVPAYGATGSGKTYTMQGTEDFPLG 124
QY 123 IPRALMDLLQLTREGAAGRPWALSVMYSLEYIQEYKVLDDLDPASGDVIREDCRGNIL 182
D 125 IPLAASITILACT-----GTWC-SVLSIYEVYEMERCYDLLEPKAKEIMALDDKGNQ 177
QY 183 IPGLSQPISFADFPERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFRQ 242

```

Db 118 LKXSTVLALCTGT-----WC-SVEISYEVYMERVCYDILLBPKAREIMVLDDKGNL 167
Qy 182 LIPGLSOKPISSFADFERHFLPASRNRTVGTATLNQSSRSASHAVLLVKVDQORERLAPRQ 241
Db 168 QKGLAWPVRSLEFHEIYSIGVQRKVAHTGLNDVSSRSASHAVLSIRITT-----DV 220
Qy 242 REGKLYLI---DIAGSEDRRTNGKGLRLKESGAINTSLFVLGKVDALNOGLPRVPYRD 298
Db 221 VKGKLNLTSDLAGNEDNRTNCGIRLQESAKINOSLFSALNVIKPEPRIPYRE 280
Qy 299 SKLTRLQDSLGSGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPTNESLOPHA 358
Db 281 SKLTRLQDSLGSGSHAVMIACLNPVE--YOEAVHTVSLAARSHRVN-----HM 328
Qy 359 LGPVKLSQKELLGPPEAK-----RARGPEE-----EIGSPPEPMAAPASQ 400
Db 329 SSASKXKDKVDM---EAKLRAWLESKGKTKSIQRMWDGLLSPNAIKTFLSMSQ 377
RESULT 9
US-10-334-143-33
; Sequence 33, Application US/10334143
; Publication No. US20040009549A1
; GENERAL INFORMATION:
; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD
; FILE REFERENCE: 038602/1543
; CURRENT APPLICATION NUMBER: US/10/334,143
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: 60/343,169
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 1237
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-334-143-33

Query Match 23.4%; Score 578.5; DB 15; Length 1237;
Best Local Similarity 32.4%; Pred. No. 1e-39;
Matches 157; Conservative 97; Mismatches 172; Indels 59; Gaps 13;
Qy 20 VRVAVRLRPFDGCTAGASDPPC-----VRGMDSCSLEIANWRNHQETLKYQDAP 69
Db 13 VRVALRCRPLVPKEISGQCMLSFVPGETQVVVGTDK-----SFTYDFV 57
Qy 70 YGERSTQODIYAGSVQPIRLHLEGQNASVLAYGPTGAGKTHMTLGS-----PEQP--GV 122
Db 58 FDPCTEQEEVFNKAVAPLIKGIFKGYNAVTLAYGOTSGKTYSGGAYTAQEENETVGI 117
Qy 123 IPRALMDLLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDDPA--SGDLVIREDCRGN 180
Db 118 IPR-----VIQLLFEIKDKSDFEFTLKVSYLEIYNEEILDLCPREKAQINIREPK 173
Qy 181 ILIPLGLSOKPISSFADFERHFLPASRNRTVGTATLNQSSRSASHAVLLVKVDQORERLAPR 240
Db 174 IKIVGLTEKTVLVALDTVSCLEQGNNSRTVASTAMNSQSSRSHAFITISLEQRKKDKNC 233
Qy 241 QREKGLYLIDLAGSEDRNRTNGKGLRLKESGAINTSLFVLGKVDALNOGLPR--VPYRD 298
Db 234 SFRSKLHLVDLAGSERQKKTAEGRDLKGININRGLCLGNVISALGDDDKGSGFPYRD 293
Qy 299 SKLTRLQDSLGSGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPTNESLOPHA 358
Db 294 SKLTRLQDSLGSGSHAVMIACLNPVE--YOEAVHTVSLAARSHRVN-----HM 328
Qy 359 LGPVKLSQKELLGPPEAKRARGPEEEIIGSPPEPMAAPASQKLSPLQKLSMDPAMLER 418
Db 352 AELNHLKQ-----QVQQLQVLLQAHGTTLPFGSINAEFSEN---LQSLMEKNQSLVEE 401

Qy 419 LLSLDRLLASQSGQAPLLSTPFRERVMVLMKTVEEK-DLEIERLTKTKQ-ELEAKMLAQK 476
Db 402 NEXLSRCLSKAAGCTAQML-----ERILITEQVNEKLNKLEELRQHVACKLDLOKLVE 456
Qy 477 ABEKE 481
Db 457 LEDQE 461
RESULT 10
US-10-116-712-670
; Sequence 670, Application US/10116712
; Publication No. US20030194764A1
; GENERAL INFORMATION:
; APPLICANT: Baugur, Chaitanya S.
; APPLICANT: Switzer, Ann
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.568
; CURRENT APPLICATION NUMBER: US/10/116,712
; CURRENT FILING DATE: 2002-04-07
; NUMBER OF SEQ ID NOS: 670
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 670
; LENGTH: 1232
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-712-670
Query Match 23.2%; Score 574.5; DB 14; Length 1232;
Best Local Similarity 32.5%; Pred. No. 2.2e-39;
Matches 160; Conservative 97; Mismatches 161; Indels 75; Gaps 15;
Qy 20 VRVAVRLRP-----FVDCTAGASDPPCVRGMDSCSLEIANWRNHQETLKY 64
Db 10 VRVALRCRPLVPKEISGQCMLSFVPG-----EPQVVVGTDK-----SF 49
Qy 65 QFDIFYGERSTQODIYAGSVQPIRLHLEGQNASVLAYGPTGAGKTHMTLGS-----PEQ 119
Db 50 TYDFVDFPSTEQEEVFNKAVAPLIKGIFKGYNAVTLAYGOTSGKTYSGGAYTAQEEN 109
Qy 120 P--GVIPRALMDLLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDDPA--SGDLVIRE 175
Db 110 PTVGVIPR-----VIQLLFEIKDKSDFEFTLKVSYLEIYNEEILDLCPREKAQINIRE 165
Qy 176 DCRGNLILPGLSOKPISSFADFERHFLPASRNRTVGTATLNQSSRSASHAVLLVKVDQOR 235
Db 166 DPKEGKIVGLTEKTVLVALDTVSCLEQGNNSRTVASTAMNSQSSRSHAFITISLEQRKK 225
Qy 236 LAPFRQREGKLYLIDLAGSEDRNRTNGKGLRLKESGAINTSLFVLGKVDALNOGLPR-- 293
Db 226 SDKNSSFRSKLHLVDLAGSERQKKTAEGRDLKGININRGLCLGNVISALGDDDKGKG 285
Qy 294 VPYRDSKLTRELQDSLGSGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPTNES 353
Db 286 VPYRDSKLTRELQDSLGSGSHAVMIACLNPVE--YOEAVHTVSLAARSHRVN-----HM 328
Qy 354 LQPHALGPVK--LSQKELIGPPEAKRARGPEEEIIGSPPEPMAAPASQKLSPLQKLS 410
Db 346 PQTAELNHLKQVQQLQVLL-----LQAHGG-----TLFGSITVPESENLOSLME 390
Qy 411 MDPAMLERLLSLDRLLASQSGQAPLLSTPFRERVMVLMKTVEEK-DLEIERLTKTKQ-EL 468
Db 391 KQNSLVEEKEKLSRGLSEBAGTQAML-----ERILITEQVNEKLNKLEELRQHVACK 445
Qy 469 EAKMLAQKAEKE 481
Db 446 DLQKLVETLEDQE 458
RESULT 11
US-10-408-765A-2153
; Sequence 2153, Application US/10408765A

LENGTH: 776

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_101324C.1.pap

US-10-437-963-103943

Query Match 23.1%; Score 571; DB 16; Length 776;

Best Local Similarity 31.5%; Pred. No. 2.2e-39;

Matches 152; Conservative 99; Mismatches 194; Indels 38; Gaps 10;

QY 18 ARVAVAVLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQETLK-----YQDAP 69

DB 204 SRIMVFLRPMRSRKEKDGASRSCVKIYNKDDVLTPEFASSETDYLRLKVRGRHFCDSS 263

QY 70 YGERSTQDIIYAGSVQPIRLHLEGGONASVLAYGPTGAGKTHMLGSPQPGVTPRALMD 129

DB 264 FPDITTAQEVYSTTSDIVEGVQGRNGTVFCYATGAGKTYTMLGTMSFGVWVLAIKD 323

QY 130 LLQLTREGAERGPWALSVMYSLEYIYQEKVLDLDPASGDLVIREDCRGNILIPGLSQK 189

DB 324 LFTKVRQRSHDGNH---SIQLSYLEVYNETVRDLSLPCR-FULLREDKQGV-AAGLTHY 378

QY 190 PISSEFADPERFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQRE-RLAPPRQEGKYL 248

DB 379 RAYSTDVEMKLLQGNQNRTEPTRVNETSSRSHAILQVIVVEYSIDGGSIVTRVGKLSL 438

QY 249 IDLAGSEDRNRRTGNKGLRKESGAINTSLSFLVGKVDALNOGLPRVPYRDSKLTLLQDS 308

DB 439 IDLAGSERALATDQRTQSIIEGANINRSLALSSCINALVEGKKHPIYRNSKLTLLQDS 498

QY 309 LGGSAHSILIANIAPERFYLDTVSALNFAARSKEVINRPTNESLOPHALGPVKLSQKE 368

DB 499 LGGSCNTVMIANISPSNLSFGETQNTLHWADRAKEIKTKALTANEVRLVTDSETQAK 558

QY 369 LL-----GPPAKRARGPEEEIGSPPEMAAPASAKLSPLQKLSNDPAMLERLLSLD 423

DB 559 LVLELOKENSELRQQLAQKLLTVQAQTLASNASPQSPAPSAQISTPCSTQKVKRS 618

QY 424 RLLASQSGAPLLSTPKRERV-----LMTVEEKLEIERLTKKELEAKMLAQ 475

DB 619 ILAGN-----CFNTPDSKRAAENAAQVRDLQRKVKAMEABIEKMK-KEHLLQK--Q 667

QY 476 KAE 478

DB 668 KDE 670

RESULT 14

US-10-425-115-332919

Sequence 332919, Application US/10425115

Publication No. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Plants

TITLE OF INVENTION: Plants

FILE REFERENCE: 38-21(53222)B

CURRENT APPLICATION NUMBER: US/10/425,115

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 332919

LENGTH: 783

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

NAME/KEY: unsure

LOCATION: (1)...(783)

OTHER INFORMATION: unsure at all xaa locations

FEATURE:

OTHER INFORMATION: Clone ID: MRT4577_66737C.1.pap

US-10-425-115-332919

Query Match

Best Local Similarity 23.1%; Score 571; DB 17; Length 783;

Matches 163; Conservative 91; Mismatches 195; Indels 36; Gaps 12;

QY 8 KIGA-TRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQETLK--- 63

DB 197 EVGAGTAAEAESRILVFLRPMRSRKEAGSRSCVKIYNKRVFLTESSENDVYLRUKR 256

QY 64 -----YQDAPYGERSTQDIIYAGSVQPIRLHLEGGONASVLAYGPTGAGKTHMLGSP 118

DB 257 GRDSHFCDSPFPDSTTQAEVYSTADLVEGVQGRNGTVFCYATGAGKTYTMLGTME 316

QY 119 QPGVIPRALMDLLQLTREGAERGPWALSVMYSLEYIYQEKVLDLDPASGDLVIREDCR 178

DB 317 NPGVWVLAINDLFSKVTKQKH-----SIKLSLEYIYNETVRDLSLPCR-FULLREDKQ 368

QY 179 GNILIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQRELRAP 238

DB 369 G-IVAAGLTQRSVYSTDEVMELLQKGNKRTTEPTRVNETSSRSHAVLQVVVEYRSLDGV 427

QY 239 PRQREGKLYLDLAGSEDRNRRTGNKGLRKESGAINTSLSFLVGKVDALNOGLPRVPYR 298

DB 428 NVKAGAGKSLIDLAGSERALATDQRTQSIIEGANINRSLALSSCINALVEGKKHPIYRN 487

QY 299 SKLTRLODSIGGSAHSILIANIAPERFYLDTVSALNFAARSKEV---INRPTNESL- 354

DB 488 SKLTQLLKSLGGACNTVMIANISPSNLSFGETQNTLHWADRAKEIKTKOQTTVNEEVL 547

QY 355 -QPHALGPVKLS-QKE--LLGPPPEAKRARGPEEEIGSPPEMAAPASAKLSPLQKLS 410

DB 548 DQDSETMVLVELOKENVRLREQLAK-----QQKLLTAAQLLASKTSPPSPAPSHV 602

QY 411 MDPAMLERLLSLDRLLASQSGAPLLS-----TPKRERMVLMKTVEEKDLIERLTKQK 466

DB 603 STPGSTQRTKRRSILAAAGGNYFSLDSKRHAADNAQVRELQKRVSTLSEIEKMK-KEH 661

QY 467 ELEAK 471

DB 662 LLQIK 666

RESULT 15

US-10-425-114-59708

Sequence 59708, Application US/10425114

Publication No. US20040034889A1

GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E.

APPLICANT: Tabaska, Jack E.

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53313)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 59708

LENGTH: 854

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: LIB3689-256-H1_FLI.pap

US-10-425-114-59708

Query Match

Best Local Similarity 23.1%; Score 571; DB 15; Length 854;

Matches 163; Conservative 91; Mismatches 195; Indels 36; Gaps 12;

QY 8 KIGA-TRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQETLK--- 63

```

Db      268 EVGAGTAEAESRILVFLRPMRSKEKEAGSRSCVKIVNRKEVFLTESASENDYLRKR 327
QY      64 -----YQDAFYGERSTQDIYAGSVQPIRLHLLBQGNASVLA YGPTGAGKTHMLGSP 118
Db      328 GRDSHFCFDSVFPDSTTQAEVYSTADLVEGVLOGRNGTFCYCATGAGKTYTMLGTME 387
QY      119 QGVIPRALMDLLOLTREBAGRCPWALSVTMSYLEIYQEKYLDLDDPASGDLVIREDCR 178
Db      388 NFGVMVLAINDLFSKVTQKNH-----SILSYLEIYNETVRDLSLPGS-PLNLRDQK 439
QY      179 GNILIPGLSQKPTSSFPADPERHFLPASRNRITVGATRLNQSRSHAVLIVKVDQERLAP 238
Db      440 G-IVAAGLTQRSVYSTDEVWELLQKGNKRTTEPTFVNETSSRSHAVLQVVVEYRSLDGV 498
QY      239 FRQREGKLYLIDLAGEDNRTGNKGLRKESGAINTSFLVIGKVVDALNQGLPRVPYRD 298
Db      499 NVKRACKLSLIDLAGSERALATDQRTQRTGIEGANINRSLALSSCINALVEGKCHIPYRN 558
QY      299 SKLTRLLQDSLGSASHSILIANIAPERRFYLDITVSALNFAARKEV---INRPFTNESL- 354
Db      559 SKLTQLLQDSLGSACNTVMIANISPSNLSFGETQNTLHWADRAKEIKTKTOOTVNEEVL 618
QY      355 -QPHALGPVKLS-QKE--LIGPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSS 410
Db      619 DQPDSETMLVLELQENRVLREQLAK-----QQKLLTAEAOQLLTSKTSQPSAPPSPHV 673
QY      411 MDPAMLERLLSLDRLLASOGSOGAPLLS-----TPKRRMVLMTVEEKDLEIERLKTOK 466
Db      674 STPGSTQRTRRSILAAGGNYFSKLDKSRHAADNAQVRELQKVKSTLSEIEKMK-KEH 732
QY      467 ELEAK 471
Db      733 LLOLK 737

```

Search completed: November 5, 2004, 19:09:49
 Job time : 83.6338 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 ; Search time 14.1983 Seconds
(without alignments)
3300.235 Million cell updates/sec

Title: US-10-797-893-8
Perfect score: 2472
Sequence: 1 MGRCLSKIGATRRPPPARV.....LEAKVLAQKAEKENHCPMT 487

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2014	81.5	665	2	S62328
2	614	24.8	664	2	T48258
3	612	24.8	784	1	A55236
4	597.5	24.2	703	2	A86319
5	579.5	23.4	1226	2	T51617
6	567	22.9	1254	2	T18277
7	566	22.9	1225	2	A56514
8	565	22.9	1058	2	T47525
9	563.5	22.8	701	1	B44259
10	563	22.8	747	1	A57107
11	561.5	22.7	786	2	A53939
12	561.5	22.7	1056	2	H84777
13	560.5	22.7	1006	2	T02017
14	560.5	22.7	1076	2	B84687
15	560	22.7	885	2	D86151
16	558	22.6	1231	2	A54803
17	557.5	22.6	1388	2	T30335
18	556.5	22.5	1066	1	A48669
19	556	22.5	699	1	A38982
20	551	22.3	2954	2	T14156
21	545	22.0	1229	2	T748929
22	544.5	22.0	935	2	T51930
23	543	22.0	928	2	T10164
24	542	21.9	1070	2	T06733
25	542	21.9	1695	2	A56921
26	541.5	21.9	784	2	T40594
27	541.5	21.9	793	2	JC5831
28	540.5	21.9	929	2	T51932
29	536	21.7	1022	2	E84792

ALIGNMENTS

RESULT 1

S62328
kinesin-like DNA binding protein KID - human
C:Species: Homo sapiens (man)
C>Date: 23-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 16-Aug-2004
C:Accession: S62328
R:Tokai, N.; Fujimoto-Nishiyama, A.; Toyoshima, Y.; Yonemura, S.; Tsukita, S.; Inoue, .
EMBO J. 15, 457-467, 1996
A>Title: Kid, a novel kinesin-like DNA binding protein, is localized to chromosomes an
A:Reference number: S62328; MUID:96174806; PMID:8599929
A:Accession: S62328
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-665 <TOK>
C:Superfamily: kinesin motor domain homology
C:Keywords: ATP; nucleotide binding; P-loop
F:44-374/Domain: kinesin motor domain homology <KMOF>
F:128-135/Region: nucleotide-binding motif A (P-loop)

Query Match 81.5%; Score 2014; DB 2; Length 665;
Best Local Similarity 82.4%; Pred. No. 1e-128;
Matches 416; Conservative 3; Mismatches 44; Indels 42; Gaps 4;

Qy	2	GRCLSKIGATRRPPPARVAVRLRPVDCGTAGASDPPCVRGWMDSCSLEIANWRNHQET	61
Db	26	GRCLSKIGATRRPPPARVAVRLRPVDCGTAGASDPPCVRGWMDSCSLEIANWRNHQET	85
Qy	62	LKYQDAFYGERSTQDDIYAGSVQPIRLHLLLEGQNAS-VLAYGPTGAGKTHTMLGSPEOP	120
Db	86	LKYQDAFYGERSTQDDIYAGSVQPIRLHLLLEGQNAKVVLAYGPTGAGKT-THAQGPRA	144
Qy	121	GVIPRALMDLLQLTRREGAQRWALSVTMSYLSIYQKVLDDLDPAGDGLVIREDCRGN	180
Db	145	WGDPAGSHGPPAAHKGGCGRPGAMGLSVTMSYLSIYQKVLDDLDPAGDGLVIREDCRGN	204
Qy	181	ILIPGLSQKPTSSPADFERHFLPASRNTVCATRLNQRSSRSHAVLLVKVQDRELAPEP	240
Db	205	ILIPGLSQKPLISSPADFERHFLPASRNTVCATRLNQRSSRSHAVLLVKVQDRELAPEP	264
Qy	241	QREGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFVLGVKVDALNQGLPRVYRDSK	300
Db	265	QREGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFALGVKVDALNQGLPRVYRDSK	324
Qy	301	LTRLQDSLGSSAHSILIANIAPERRFVLTVSALNFAARKEVINRFTNESLOPHALG	360
Db	325	LTRLQDSLGSSAHSILIANIAPERRFVLTVSALNFAARKEVINRFTNESLOPHALG	384
Qy	361	PVKLSQKELGPPPAKARGPEEEIGSPPEMAA------PASASQ	400
Db	385	PVKLSQKELGPPPAKARGPEEEIGSPPEMASSLCCLPTQPTPEAKAAWTRPCGAPP	444
Qy	401	KLSPLOKLSSMDPAMLERLLSLDLLASQSQGAPLLSTPKRERMVLMKTVEEKDLEIR	460

Db 445 QLGPSACL-----PGEPCALLSTPKRVMVLMKTVEEKDLEIER 484

QY 461 LXTKKELEAKMLAKAEKENHCP 485

Db 485 LXTKKELEAKMLAKAEKRTIVP 509

RESULT 2

T48258

kinesin-like protein - Arabidopsis thaliana

N;Alternate names: protein T1E22.130

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004

C;Accession: T48258

C;Revan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft, submitted to the Protein Sequence Database, March 2000

A;Reference number: 224489

A;Accession: T48258

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-664 <REV>

A;Cross-references: UNIPROT:Q9L288; EMBL:AL162874

A;Map position: 5

A;Experimental source: cultivar Columbia; BAC clone T1E22

C;Genetics:

A;Description: may be part of a motor protein that provides anterograde fast axonal tra

A;Superfamily: kinesin-related protein Kif3; kinesin motor domain homology

C;Keywords: ATP; coiled coil; microtubule binding; nucleotide binding; P-loop

F;10-349/Domain: head globular #status predicted <HGL>

F;120-350/Domain: kinesin motor domain homology <KMT>

F;106-113/Region: nucleotide-binding motif A (P-loop)

F;350-580/Domain: helical rod #status predicted <ROD>

F;581-784/Domain: tail globular #status predicted <TGL>

F;112/Binding site: ATP (lys) #status predicted

Query Match 24.8%; Score 614; DB 2; Length 664;

Best Local Similarity 32.4%; Pred. No. 8.2e-34; Indels 64; Gaps 13;

Matches 163; Conservative 94; Mismatches 187;

QY 18 ARVRVAVRLRPFFV-----DGTAGADPPPCVGRGMDSCSLFIANRHNHQLK---YQFDA 69

Db 78 SNRVVLVRVLPFLPREISDSCDGRSCVSVIGDGGDTSEVAVYKDPDSCRNSYQLDA 137

QY 69 FYG-ERSTQDIYAGSVQPIRLHLEGQNASVLAYGPTGAGKTHMLGSPQGVIPRAL 127

Db 138 FYGREDNNVXHFIDREVSPLPIGFHGFNATVAYGATGSGKFTTQGIDELPLMLPTM 197

QY 128 MDLLQLTREAGRPMWLSVYIYQKVLMDLPASGDLVIREDCRGNILIPGLS 187

Db 198 STILSMCEKTSRAE-----ISYEVYMDRCWDLLEVKDNEIAVWDDKQGVHLKGLS 250

QY 188 QKPISSPADFERHFLPASRNTVGATRLNQRSSSHAVLVKVQDRLAPFRQEGKLY 247

Db 251 SVPVKSMSEFOEAYLCGVQRKVAHTGLNDVSSRSHGVLSVTSQGLVT-----GKIN 304

QY 248 LIDLAGEDNRTGNKGLRKESGAINTSFLVLGVVDALNOGLPRVPYRDSKLTRLLQD 307

Db 305 LIDLAGEDNRTGNKGLRKESGAINTSFLVLGVVDALNOGLPRVPYRDSKLTRLLQD 364

QY 308 SLGGSASHILTIANTAPERFRFYDTVSALNFAARKEVINRPFVTNIESLQPHALGP-VKLS- 365

Db 365 SLGGSASHILTIANTAPERFRFYDTVSALNFAARKEVINRPFVTNIESLQPHALGP-VKLS- 418

QY 366 QKELLGPPEAKRGPPEEETGSPPEPMAAPASAKLSPLQKLSMDPAMLE-----RVLMKMTVEEK 454

Db 419 EAKLQAWLESKGMKSAHRMAIRSPMLGTNQTISQSSVVKLLCHRSATAESAKLAGTG 478

QY 418 -----RLSLDRLLASQSGAPL-----LSTPKRE-----RVLMKMTVEEK 454

Db 479 QRDAPVTARNLFGVETTLAASH--LWEPTRNLQASPTKEDRDTSGEENLLVSEASLRDN 536

QY 455 DLEIRLKTQKKELEAKMLAKAEKEN 482

Db 537 TLDVEKKYTELSPLEALSPIDSNAKPN 564

RESULT 3

A55236

kinesin-related protein KLP68D - fruit fly (Drosophila melanogaster)

N;Alternate names: kinesin-like protein 5; KLP5

C;Species: Drosophila melanogaster

C;Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 09-Jul-2004

C;Accession: A55236; E41298

R;Pesavento, P.A.; Stewart, R.J.; Goldstein, L.S.B.

J;Cell Biol. 127, 1041-1048, 1994

A;Title: Characterization of the KLP68D kinesin-like protein in Drosophila: possible ro

A;Reference number: A55236; MUID:95050960; PMID:7525600

A;Accession: A55236

A;Molecule type: mRNA

A;Residues: 1-784 <PES>

A;Cross-references: UNIPROT:P46867; GB:U15974; NID:9595912; PIDN:AAA69929.1; PID:g56509

R;Stewart, R.J.; Pesavento, P.A.; Woerpel, D.N.; Goldstein, L.S.B.

Proc. Natl. Acad. Sci. U.S.A. 88, 8470-8474, 1991

A;Title: Identification and partial characterization of six members of the kinesin supe

A;Reference number: A41298; MUID:92020874; PMID:1924306

A;Accession: E41298

A;Molecule type: DNA

A;Residues: 'TC', 222-337, 'VRGOV' <STE>

A;Cross-references: GB:M74431; NID:g157791; PIDN:AAA28658.1; PID:g157792

C;Genetics:

A;Gene: FlyBase:Klp68D; KLP5

A;Cross-references: FlyBase:FBgn0004381

C;Function:

A;Description: may be part of a motor protein that provides anterograde fast axonal tra

C;Superfamily: kinesin-related protein Kif3; kinesin motor domain homology

C;Keywords: ATP; coiled coil; microtubule binding; nucleotide binding; P-loop

F;10-349/Domain: head globular #status predicted <HGL>

F;120-350/Domain: kinesin motor domain homology <KMT>

F;106-113/Region: nucleotide-binding motif A (P-loop)

F;350-580/Domain: helical rod #status predicted <ROD>

F;581-784/Domain: tail globular #status predicted <TGL>

F;112/Binding site: ATP (lys) #status predicted

Query Match 24.8%; Score 612; DB 1; Length 784;

Best Local Similarity 33.0%; Pred. No. 1.4e-33; Indels 88; Gaps 18;

Matches 175; Conservative 84; Mismatches 184;

QY 12 TRRP-----PPARVRVAVRLRPFFVDGTAGADPPCV-----RGMDCSLFIANW--RN 57

Db 5 SRPPTGSSQTNPCVQVVRCPMSNRBSERSPEVVNVVYVNRGV----VELQNVDGN 60

QY 58 HOETLYKQPDAPYGRSTOODIYAGSVQPIRLHLEGQNASVLAYGPTGAGKTHM---L 114

Db 61 KEQRKVFTYDAAYDASATQTTLYHEVFPVLSVLEGGNGCIFAYGQTGTGKTFTMEGVR 120

QY 115 GSPEQGVIPRALMDL-LQLTREAGRPMWLSVYIYQKVLMDLPASGDLVI 173

Db 121 GNDELGGIIPRTFEQIWLHINTEN-----FQFLVDVSVLEIYMBELRDLKPKNSKHLEV 175

QY 174 REDCRGN-LIIPGLSQKPISSPADFERHFLPASRNTVGATRLNQRSSSHAVLVKVQDQ 232

Db 176 RE--RSGVYVYVNLHAINCKSVEDMIKVMQVGNKRTVGTNNNEHSSRSHAFIMIKIEM 233

QY 233 RELAPFRQEGKLYLIDLAGEDNRTGNKGLRKESGAINTSFLVLGVVDALNOGLP 292

Db 234 CD-TETNTIKVGNLIDLAGESRQSKTGASERLKEASKINLALSLSLGNVISBALAESPP 292

QY 293 RVPYRDSKLTRLLQDSLGGSASHILTIANTAPERFRFYDTVSALNFAARKEVINRPFVTNE 352

Db 293 HVPYRDSKLTRLLQDSLGGSASHILTIANTAPERFRFYDTVSALNFAARKEVINRPFVTNE 352

QY 353 SLOPHALGPVKLSQ-----KELGPPPEAKA-----RGPEEEI----- 386

Db 353 DPQ-----DAKLEYQBEIERLKLKLPQOQQRSEKQVTAQRKVRKPKKETTVMKMSDS 407

QY 387 -----GSPEPMAAPASAKLSPLQKLSMDPAMLERLLSLDRLLASQSGQ 433

Db 408 LQVSTTBQPVEDSDPE---GAESSEKNEAEVAKNSNEELERVEN-SKLAALKAELE 463

QY 434 APILSTPKRERMVLMKTVEEKDLEIER-----LTKKKELEAKMLAKAE 479

Db 464 GQVVRGKN---LLDYSERQIELEKLVIEAKRKRREIQQQLQEQE 510

RESULT 4

A86319
F25116.11 protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: A86319
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408: 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A86319
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-703 <STO>
A:Cross-references: UNIPROT:Q9FZ77; GB:AE005172; NID:g9795601; PIDN:AAF98419.1; GSPDB:GN
C:Genetics:
A:Map position: 1

Query Match 24.2%; Score 597.5; DB 2; Length 703;
Best Local Similarity 33.9%; Pred. No. 1.2e-32;
Matches 173; Conservative 88; Mismatches 172; Indels 77; Gaps 17;
Qy 18 ARVAVRLRPFDVGTAGADPPCVRGMDSCSLSIANRNHQTLLK-----YQDAF 69
Db 150 SRLVFLRPMGKRENGSRCCVKLNKRDVLTFTENDYLRKLRLVRHFTDSS 209
Qy 70 YGERSTQDIYAGSVQPIRLHLLLEGONASVLAYGPTGAGKTHMLGSPQGVTPRALMD 129
Db 210 FPETTTQGVYSTTTGDLVEAVLEGNGSVFCYCATGAGKTYTMTGNTWENPVMVLAID 269
Qy 130 LLQLTREGAEGRPWALSVTMSYLEIYQEKVLDLLDPAAGDLVREDRCGNILIPGLSOK 189
Db 270 LFAVVRSLDGNH---VHLSYLEVNETVRDLSLSPOR-PLIUREDKQ-----VMALLQR 321
Qy 190 PISFADPFERHFLPASNRRTVGATRLNQSRSHAVLLVKVDQRLERLAPFR--OREGKLY 247
Db 322 -----GNQNRITTEPRCNETSSRSRAILQVIVEYKTRDASNNIISRVGKLS 367
Qy 248 LIDLAGSDNRRTGNKGLRKESGAINTSFLVLGKVVDALNQGLPRVPYRDSKLTLLQD 307
Db 368 LIDLAGSERALATDQRTLSLEGANINRSLIALSSCINALVEGKKHPIYRNSKLTQLLKD 427
Qy 308 SLGGSASHILIANIAPERRFLDTVSALNFAARKEV-INRPFTNESILOPHALGPVKLSQ 366
Db 428 SLGGSCTVMIANSPSSQSGFQTNTLHWADRAKEIRVKEVEVNEEV-----VQVGE 480
Qy 367 KELGPPPEAKRARGPEEEIGSPFPMAAPASAKLSPLQKLS-----SMDPAML 416
Db 481 EE--GAQAKLLLELQXE---NSELRVQLAKQQQKLLTLQAEATAANNNNISLTTPSI 535
Qy 417 ERLLSLDRLLASQSGQA--PLLS-----TP-----KRRMV--LMKTVEKDLIERLKL 462
Db 536 SSLMTPTPSALTAQKKPRHSLLSGCTPTPESLARTKAEAEVAKELQTLVRLKALMEMERMK 595
Qy 463 -----TKOK-ELEAKWLAQAEKENHC 484
Db 596 REHGLQMKQKQDLMDLCLSRKSKTPERC 625

RESULT 5

I51617
kinesin-like protein 1 - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Aug-2004

C:Accession: I51617; A48835; S48837
R:Vernos, I.; Raats, J.; Hirano, T.; Heasman, J.; Karsenti, E.; Wylie, C.
Cell 81, 117-127, 1995
A:Title: Xklp1, a chromosomal Xenopus kinesin-like protein essential for spindle organ-
A:Reference number: A56221; MUID:95236444; PMID:7720067
A:Accession: I51617
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1226 <VER>
A:Cross-references: UNIPROT:Q91784; EMBL:X82012; NID:g562792; PIDN:CAA57639.1; PID:g56-
R:Vernos, I.; Heasman, J.; Wylie, C.
Dev. Biol. 157, 232-239, 1993
A:Title: Multiple kinesin-like transcripts in Xenopus oocytes.
A:Reference number: A48835; MUID:93246065; PMID:8482413
A:Accession: A48835
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 9-162, 'L', 164-338 <VE2>
A:Experimental source: oocyte
A:Note: sequence extracted from NCBI backbone (NCBIP:130975)
C:Genetics:
A:Gene: klp1
C:Superfamily: kinesin motor domain homology
C:Keywords: ATP; nucleotide binding; P-loop
F:9-343/Domain: kinesin motor domain homology <RMOT>
F:87-94/Region: nucleotide-binding motif A (P-loop)

Query Match 23.4%; Score 579.5; DB 2; Length 1226;
Best Local Similarity 32.1%; Pred. No. 4.3e-31;
Matches 157; Conservative 95; Mismatches 172; Indels 65; Gaps 13;
Qy 20 VRVAVRLRPFDVGTAGADPPCVRGMDSCSLSIANRNHQTLLK-----KYOFDAFYGRS 74
Db 9 VRVALRCPLVP-----KENNEGCKMLTFVPEEQVIVGTEKSFYDYVDFDSA 58
Qy 75 TQDIYAGSVQPIRLHLLLEGONASVLAYGPTGAGKTHMLGSPQGVTPRALMD 127
Db 59 EQEVNYSVAAPLTKGLPKGNATVAYGQTGSGKTYSMGAYTHNQENETVGVIPRTV 118
Qy 128 MDLLQLTREGAEGRPWALSVTMSYLEIYQEKVLDLLDPAAGDLVREDRCGNIL 183
Db 119 IALFREIHQ-----RPEWENLKVSYLEIYNEEILDLLYAADKTNTTISIREDPKEGKI 173
Qy 184 PGLSQKPISSFADPFERHFLPASNRRTVGATRLNQSRSHAVLLVKVDQRLERLAPFRQRE 243
Db 174 CGLTERDVKTALDTLSCLEQNSRSTVASTAMNSQSSRAIPTISIEQRKEGDKNSFR 233
Qy 244 GKLYLIDLAGSDNRRTGNKGLRKESGAINTSFLVLGKVVDALNQGLPR---VPYRDSK 300
Db 234 SKLHLVDLAGSERQKTKAEGDRLKEGINSINRGLLCLGNVISALGDESCKGGFVYRDSK 293
Qy 301 LTRLLQDSIGSSAHSILIANIAPERRFLDTVSALNFAARKEVINRPFTNESILOPHALG 360
Db 294 LTRLLQDSLGGNSHLLTIACVSPADSNMEETLNTLYADRAKIKNPKIVNTDPQAALQ 353
Qy 361 PVKLSQKELLGPPPEAKRARGPEEEIGSPFPMAAPASAKLSPLQKLSMDPAMLERLL 420
Db 354 RLKLQVQL--QVLLQAHGGLTPVNSMEPSNQLSLMERNKXLEKNG-----401
Qy 421 SLDRLLASQSGQAAPLLSTPKRERMVLMKTVEKDLIERLKTOKELE-----AKM 472
Db 402 KLSRELGEAAVQTAQFL-----EKIIMTEQONEK-----LGSQWELKQHAACKVNLQR 450
Qy 473 LAQKAEKE 481
Db 451 LVETLEDOE 459

RESULT 6

T18277
kinesin heavy chain - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

Fri Nov 12 12:26:29 2004

C:Accession: T18277
R:deHostos, B.L.; McCaffrey, G.; Vale, R.D.
submitted to the EMBL Data Library, November 1995
A:Reference number: Z18853
A:Accession: T18277
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1254 <DEH>
A:Cross-references: UNIPROT:Q94463; EMBL:U41289; NID:g1526990; PID:g1526991; PIDN:AA8077

Query Match 22.9%; Score 567; DB 2; Length 1254;
Best Local Similarity 36.1%; Pred. No. 3.1e-30;
Matches 126; Conservative 79; Mismatches 132; Indels 12; Gaps 5;

QY 15 PPARVAVLRPPFVDTAGASDPPCVRGMDSCSLEI-ANWRNHQETLKYQDAFYGER 73
DB PVSINRVVCRVPLTELEKGRNEHSIVHFDSKISIRANGP-----QFTDFRIFGVQ 77

QY 74 STQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMTLMSPEQPGVPRALMDL--- 130
DB ETQSQIFEDVAEPIVNDLGDVGHGTIIAYGQTASGKFTTWVGDPSHGIIIPRVIESIFVG 137

QY 131 LQUTREGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGNILIPGLSQKP 190
DB ISKREKDT-SLSLAFCLKISALELYNEKLYDYIASKNLNIREHKQNGIYVEGISEIV 196

QY 191 ISSFADPERHFLPASRNRTVGATRLNORSRSHAVLLVVDQDORERLAPRQREGKLYLID 250
DB ITSIEAYNPLNINNNRAIATKMSAASRSHSVLMELSQ-QNLSMESSKISKLFVD 255

QY 251 LAGSDNRRTGKGLRKSGAINTSLFVGLKVVDMALNOGLPRVPRYRDSKLTRELQDSLG 310
DB LAGSDISHTGAEGRMQEAKNINLSALGKIVINALTCGVNVPYRDSKLTRELQDSLG 315

QY 311 GSAHSILIANIAPERRFLDTVGSALNFAARKEVINRPTNESLQPHAL 359
DB GNSKTSLIINCSNNNEHETITLQFGTRAKTIKQPKINKKITYHEL 364

RESULT 7
A56514
Chromokinesin - chicken
C:Species: Gallus gallus (chicken)
C:Date: 21-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 09-Jul-2004
C:Accession: A56514; I50691
R:Wang, S.Z.; Adler, R.
J. Cell Biol. 128, 761-768, 1995
A:Title: Chromokinesin: a DNA-binding, kinesin-like nuclear protein.
A:Reference number: A56514; MUID:95181533; PMID:7876303
A:Accession: A56514
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1225 <WAN>
A:Cross-references: UNIPROT:Q90640; GB:U18309; NID:g603760; PIDN:AAC59666.1; PID:g603761
R:Wang, S.Z.; Adler, R. U.S.A. 91, 1351-1355, 1994
Proc. Natl. Acad. Sci.
A:Title: A developmentally regulated basic-leucine zipper-like gene and its expression in F-11-344/Region: kinase motor domain homology <KMT>
A:Reference number: A53451; MUID:94151328; PMID:8108415
A:Accession: I50691
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 728-1086, 'RI' <WA2>
A:Cross-references: EMBL:U04821; NID:g440792; PIDN:AA18960.1; PID:g440793
C:Genetics:
A:Gene: sw3-3
C:Superfamily: unassigned kinesin-related proteins; kinesin motor domain homology
C:Keywords: ATP; DNA binding; mitosis; nucleotide binding; P-loop
F:11-344/Region: kinase motor domain homology <KMT>
F:88-95/Region: nucleotide-binding motif A (P-loop)

Query Match 22.9%; Score 566; DB 2; Length 1225;
Best Local Similarity 31.3%; Pred. No. 3.5e-30;
Matches 159; Conservative 103; Mismatches 182; Indels 64; Gaps 14;

QY 17 PARVAVLRPPFV-DGTAGA-----SDPPCVRGMDSCSLEIANWRNHQETLKYQF 66
DB PVRV---VRCRPLVPKETSEGCQCLSFVPGEQVIVGSDKA-----FTY 51

QY 67 DAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMTLMS-----PEQ 119
DB DYVDFDSVEGEVFNATVAPLIRGIFGYNATVAYGOTSGKTYSMGGTYTASQEHDP 111

QY 120 PGVPRALMDLLOLTREGEAEGRPWALSVTMSYLEIYQEKVLDLDPDPA---SGDLVIRE 176.
DB MGVIPT---VIKLLFEKEQDQWFEVLKVSYLEIYNEDILDLCCSRSSQISIRE 167

QY 177 CRGNILIPGLSQKPISSPADFERHFLPASRNRTVGATRLNORSRSHAVLLVVDQDORERL 236
DB PKEGIKIVGLTERNVARSADTVSCLEQGNCRVTASTAMNSQSSKSHAFICIDQKK 227

QY 237 APPRQREGKLYLIDLAGSDNRRTGKGLRKSGAINTSLFVGLKVVDMALNOGLPR--- 293
DB DKNSSFHKLHLVDLAGSERQKTKAEGDLKEGININRGLLCLGNVISALGEENKKG 287

QY 294 VYRDSKLTRELQDSLGSAHSILIANIAPERRFLDTVGSALNFAARKEVINRPTNES 353
DB VYRDSKLTRELQDSLGSGNSHTMIACVSPADSNLEETLNTLYADRAKIKKPKIV 347

QY 354 LQPHALGPVK--LSQKELL-----GPPEAKRARGPEEEIGSPPEMAAPASAK 403
DB POAAELNHLKQVQVQLVLLQAHGGTLPVSNWAPSENLOSLMEKQSLMEENEKLS 407

QY 404 PQKLSMDPAMLERLLSLDRLLASQSGQAPLL---STPKRERMVLMKTVEKDL 459
DB GLSEAAQTQAKMLERIIIVTEQENKNAKLEQLQQAHAVCKLDLQKLETVDEEL 467

QY 460 RLKTKQEL-----BAKMLAQKAEKEN 482
DB VIRNLQVLAQFQSESAAAAAATAEMAN 495

RESULT 8
T47525
Kinesin-related protein-like - Arabidopsis thaliana
N:Alternate names: protein F16L2.60
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C:Accession: T47525
R:Jordan, N.; Bangert, S.; Wiedemann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Rudd, S.; submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24468
A:Accession: T47525
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1058 <JOR>
A:Cross-references: UNIPROT:Q9LZU5; EMBL:AL162459
A:Experimental source: cultivar Columbia; BAC clone F16L2
C:Genetics:
A:Map position: 3
A:Introns: 56/2; 99/3; 151/3; 310/3; 343/2; 395/3; 419/1; 444/3; 462/3; 489/3; 527/1; 5
A:Note: F16L2.60
C:Superfamily: kinesin-related protein KIP1; kinesin motor domain homology.

Query Match 22.9%; Score 565; DB 2; Length 1058;
Best Local Similarity 29.5%; Pred. No. 3.4e-30;
Matches 156; Conservative 99; Mismatches 201; Indels 72; Gaps 12;

QY 5 RLKIGATRRPPPAR---VYRVAVLRPPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61
DB RSSESNTNRNDEKGVNQQVILRCPLSEDEARIHTPVVI-----SCN-----ENREV 80

QY 62 LKQY-----FDAPYGRSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGK 109
DB AATQSTAGKHIDRFADFVKVFPASQKQDLYDQAIPIVFEVLEGNCTIFAYGQTGK 140

QY 110 THTMLGS-----PEQPGVPRALMDLLOLTREGEAEGRPWALSVTMSYLEIYQEKVL 161

Db 141 TYTMEGGARKKNGEFPDAGVPIRAVKQIFDLEAQAGAE-----YSMKVTFLELYNEIS 195
 Qy 162 DLLDP-----ASGDLVIREDCRGNILIPGLSOKPISSFADFERHFLPASRRNTVG 211
 Db 196 DLLAPEETIKFVDEKSKSIAMMEDGKSVFVRGLSEIEIVTANEIYKILEGSAKRTA 255
 Qy 212 ATRLNQRSSRHAVLLVKVDQERLAPRQ--REGKLYLIDLAGSEDRNRTGNKGLRLKE 269
 Db 256 ETLNKKOSSRSHSIFSITIKENTPEGEEMIKCKKLNVLDSAGSENISRGAREGRARE 315
 Qy 270 SGANTSIFVLGKVVADALNOGLPRVYRDSKLTLLQDSLGSASHSILIANIAPERFYL 329
 Db 316 AGEINKSLTUGRVINALVHSGHIPYRDSKLTLLRESLGGTKTKTCVIATISPSIHCL 375
 Qy 330 DTVALPAARSKVINRPFNTNESLQPHALGPVKLSOKELGPPPEAKRARGPEEEIGSP 389
 Db 376 ETLSTLOYAHRANKIKKPEINQMKSVAMKDIYSEIDRL--KQEVVAAR--EKNGIYIP 432
 Qy 390 EPMAAPASAKQ--LSPLQKSSMDPAMLERLLSLDRLLASQSGGAPLLSTPKRERMVL 447
 Db 433 KDRYIQEAEKMAEKIERLELQESKDKRVVDLQELVNSQQILTAELESEKLEKTEKLL 492
 Qy 448 MK-----TVEKDLERIKTKQKELEAKMLAQKAE 478
 Db 493 ETEHSLFDLEEKYRQANATIKKEFVISNLLKSEKSLVERAFQLRTE 540

RESULT 9

B44259
 kinesin-related protein KIF3A - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 09-Jul-2004
 C:Accession: B44259; S27872
 R:Alzawa, H.; Sekine, Y.; Takemura, R.; Zhang, Z.; Nangaku, M.; Hirokawa, N.
 J. Cell Biol. 119, 1287-1296, 1992
 A>Title: Kinesin family in murine central nervous system.
 A:Reference number: A44259; MUID:93077686; PMID:1447303
 A:Accession: B44259
 A:Molecule type: mRNA
 A:Residues: 1-701 <ATZ>
 A:Cross-references: UNIPROT:P28741; EMBL:D12645; NID:G220469; PIDN:BAA02166.1; PID:G2204
 A:Experimental source: brain
 A>Note: sequence extracted from NCBI backbone (NCBIP:118911)
 C:Complex: heterodimer with KIF3B (PIR:A57107); the KIF3A/3B heterodimer associates with
 C:Function:
 A:Description: KIF3 complex is a motor protein that provides anterograde fast axonal tra
 C:Superfamily: kinesin-related protein Kif3; kinesin motor domain homology
 C:Keywords: ATP; coiled coil; heterodimer; heterotrimer; microtubule binding; nucleotide
 F:1-368/Domain: head globular #status predicted <HGL>
 F:15-351/Domain: kinesin motor domain homology <KMOT>
 F:100-107/Region: nucleotide-binding motif A (P-loop)
 F:369-599/Domain: helical rod #status predicted <ROD>
 F:600-701/Domain: tail globular #status predicted <TGL>
 F:106/Binding site: ATP (Lys) #status predicted

Query Match 22.8%; Score 563.5; DB 1; Length 701;
 Best Local Similarity 30.2%; Pred. No. 2.3e-30;
 Matches 169; Conservative 83; Mismatches 184; Indels 123; Gaps 16;
 Qy 9 IGATRRPPPA-RVRVAVRLRFP-----VDGTAGASDPPPCVRG 44
 Db 3 INKSEKESCDNVKVVRCRPLNREKSMCMVRQSVDEMGGITVHKTDSSNEPP----- 58
 Qy 45 MDSKSLIANWRNHQETLKTKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLA YGP 104
 Db 59 -----KTFTFDVFGPESKQLDVNLTARPIIDSVLEGYNGTIFAYGQ 101
 Qy 105 TGAGKTHMLGSPQEP---GVIPALMDLL-QLTREAGSRPWALSVTMSYLEIYOEKV 160
 Db 102 TGKGTPTMEGVRAVPLGRVIGVINSFAHIFGHAKAEG-----DTRFLVRVSYLEIYNEEV 157
 Qy 161 LDLL-DPASGDLVIREDCRGNILIPGLSOKPISSFADFERHFLPASRRNTVGATRLNQRS 219

Db 158 RDLGKQDQTORLEVKERPDVGVIKDLISAYVVVNADDMDRIWTMLGHKNRSVGTANNBHS 217
 Qy 220 SRSHAVLLVKVDQERLAP--FRREGKLYLIDLAGSEDRNRTGNKGLRLKESGAINISL 277
 Db 218 SRSHAITTITTECEKGVGMHVMGMKHLVLDLAGSERQAKTGATGQRLKEATKINLSL 277
 Qy 278 FVLGKVVADALNOG-LPRVPYRDSKLTLLQDSLGSASHSILIANIAPERFVLDTVSALN 336
 Db 278 STLGNVISALVDGKSTHVPYRNSKLTLLQDSLGSNTKMWCANIGPADYNYDEITSLR 337
 Qy 337 PAARSKVINRPFNTNESLQPHALGPV-----KLSQ-KELGLPPPEAKRARGPEEBEI 386
 Db 338 YANRAKNIKINAKINEDPKOALLRQFQKEIEELKKLEEGEVSGSDTSGSEEDDEGEL 397
 Qy 387 G-----SPEPMA-----APASASOKL 402
 Db 398 GEDGEKKRRDQAGKKKVSPPKVVEMQAKIDEERKALETKLDMEEEERNKARAELEER 457
 Qy 403 SPLQKSSMDPAMLERLLSLDRLLASQSGGAPLLSTPKRERMVLMKTVEEKOLIEIRLK 462
 Db 458 KDLKAQOEHQSLLEKLSALEKKVI---VGVVDLLAKAEEQE---KLEESNWELE--- 507
 Qy 463 TKQKELEAKMLAQKAEKKE 481
 Db 508 --ERRRRAEQLRKEEKE 524

RESULT 10

A57107
 kinesin-related protein KIF3B - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
 C:Accession: A57107
 R:Yamazaki, H.; Nakata, T.; Okada, Y.; Hirokawa, N.
 J. Cell Biol. 130, 1387-1399, 1995
 A>Title: KIF3A/B: a heterodimeric kinesin superfamily protein that works as a microtub
 A:Reference number: A57107; MUID:96032268; PMID:7559760
 A:Accession: A57107
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-747 <YAM>
 A:Cross-references: UNIPROT:Q61771; GB:D26077; NID:g1060922; PIDN:BAA05070.1; PID:g1060
 A:Experimental source: brain
 C:Complex: heterodimer with KIF3A (PIR:B44259); the KIF3A/3B heterodimer associates wir
 C:Function:
 A:Description: KIF3 complex is a motor protein that provides anterograde fast axonal t
 C:Superfamily: kinesin-related protein Kif3; kinesin motor domain homology
 C:Keywords: ATP; coiled coil; heterodimer; heterotrimer; microtubule binding; nucleot
 F:1-363/Domain: head globular #status predicted <HGL>
 F:10-346/Domain: kinesin motor domain homology <KMOT>
 F:96-103/Region: nucleotide-binding motif A (P-loop)
 F:364-592/Domain: helical rod #status predicted <ROD>
 F:594-747/Domain: tail globular #status predicted <TGL>
 F:102/Binding site: ATP (Lys) #status predicted

Query Match 22.8%; Score 563; DB 1; Length 747;
 Best Local Similarity 30.7%; Pred. No. 2.8e-30;
 Matches 169; Conservative 95; Mismatches 189; Indels 98; Gaps 16;
 Qy 20 VRVAVRLRPPVDGTAGASDPPPCVRGMD-----SCSLEIANWRNHQETLKTKYQDFAFYGERST 75
 Db 10 VVVVRCRP-MNGKKAASVDKVVDDVKLGQSVSKPKGTSHMPKTFDFAVDYDNNAK 68
 Qy 76 QDDIYAGSVQPIRLHLLLEGONASVLA YGPTGAGKTHTM---LGSPEQGVIPRALMDLL- 131
 Db 69 QFELYDETFRFLVDSVLQGFNGTIFAYGQTGKTGTMEGVGRGPEKRGVIPNSFDHIFT 128
 Qy 132 QLTFREAGSRPWALSVTMSYLEIYOEKVLDDLL-PASGDLVIREDCRGNILIPGLSOKP 190
 Db 129 HISRSONQ-----YLRASYLEIYQEIIRDLISKOQTKRLEKLERPDTGVYVVDLSFV 183
 Qy 191 ISSFADFERHFLPASRRNTVGATRLNQRSSSHAVLLVKVDQERLAPFRQ--REGKLYL 248

```

Db      184  TKSVEIEHVMVGNQNSVGATNMNEHSRSHAFVITIECEVGLDGENHVRGKJNL 243
QY      249  IDLAGSEDRRTNGKRLKESGAINTSFLVGLKVVDALNOG-LRPVPRYRDSKLTLLQD 307
Db      244  VDLGASERQAKTGAQERLKEATKINLSALGNVISALVDGKSTHPIYRDSKLTLLQD 303
QY      308  SIGGSAHSILIANIAPERFPYLDTVSALNPAARSKVINRPFNE----- 352
Db      304  SLGGNAKTVMANVNGPASYNVEETLTTRYANRAKNIKNKPRVNEPKDALLREFOEI 363
QY      353  ----SLQPHALGPVLSOKELL-----GPEAKRARGPEEEIG----- 387
Db      364  RLKAQLEKSGIRKKRERKREGGGGGGEBEEREGEEDGDDKDDYWRBQEKLEI 423
QY      388  -----SPEPMAAPASAKSLPQK-----LSSMDPAMLERLLSLDLLAS 428
Db      424  EKRAIVEDHSLVAERKRLKEKEKQWEDLRREKDAEMLGAKIKAMESKLLVGGKNI 483
QY      429  QSGQGAPLL-----STPKRRMV--LMKTVEKDLIER-----LKTQK-KE 467
Db      484  HTNEQQKILEQKQKEIAEQKREREIQQQMESRDEETLELKETYTSLOQOEVDIKTK 543
QY      468  LEAKMLAOKAE 478
Db      544  LFSKLQAVKAE 554

RESULT 11
A53939
kinesin homolog KHP1 - Chlamydomonas reinhardtii
C;Species: Chlamydomonas reinhardtii
C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C;Accession: A53939
R;Walther, Z.; Vashishtha, M.; Hall, J.L.
J. Cell Biol. 126, 175-188, 1994
A;Title: The Chlamydomonas FLA10 gene encodes a novel kinesin-homologous protein.
A;Reference number: A53939; MUID:94299638; PMID:8027176
A;Accession: A53939
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-786 <WAL>
A;Cross-references: UNIPROT:P46869; EMBL:L33697; NID:g497696; PID:g4976
A;Note: authors translated the codon AAC for residue 753 as Asp
C;Genetics:
A;Gene: FLA10
C;Superfamily: unassigned kinesin-related proteins; kinesin motor domain homology
C;Keywords: ATP; coiled coil; nucleotide binding; P-loop
F;11-359/Domain: kinesin motor domain homology <KNOT>
F;97-104/Region: nucleotide-binding motif A (P-loop)

Query Match      22.7%; Score 561.5; DB 2; Length 786;
Best Local Similarity 30.7%; Pred. No. 3.8e-30;
Matches 172; Conservative 88; Mismatches 190; Indels 111; Gaps 19;

QY      16  PPA-----RVRAVRLRPF-----VDGTAGASPPPCVRGMDSCSLIANWR-NHQTLK- 63
Db      2  PPAGGSGSEKVVVRCRPLNGKKGKADGRSRIVD---MDVDAGQVKVRNPKADASEPPKA 57
QY      64  YQDFAFYGRSTQDDIYAGSVQPIRLHLEGNASVLAYGPTGAGKTHMLGSPGEP--- 120
Db      58  FTFDQVYDNCQQRDVFDTARELIDSCIEGYNGTIFAYGQTGTGKSHTWEGKDEPELR 117
QY      121  GVTPRALMDLLO-LTREBAGRPWALSVMYSYLEIYQEKVLDLL-DPASGDVIREDCR 178
Db      118  GLIPNTFRVVEIIARDSGTK-----EFLVRSSYLEIYNEEVRDLGDKHKKMELKESPD 173
QY      179  GNILIPGLSQKPISSADFERHPLPASRNRTPCATRLNQESSSHAVLLVKVDQERLAP 238
Db      174  RGVYVKDLQFQVCKNYEEMKVLLAGKONRQVGAATLMOQSSRSHSIFTITIECIKLES 233
QY      239  F-----RQREGKLYLIDLAGSDNRRTGNKGLRKESGAINTSFLVGLKAVD 285

```

```

Db      234  AAAQPKAGKDDNSHVRVVKLNLVDLAGSERQDKTGADRLKEGKINLSLTALGNVIS 293
QY      286  ALNQGLP-RVPYRDSKLTLLQDSLGGSAHSILIANIAPERFPYLDTVSALNPAARSKV 344
Db      294  ALVDGKSGHIPYRDSKLTLLQDSLGNTKTVMANIGPADWNYDETMTSLRYANRAKNI 353
QY      345  INRPFTNESLQPHALGPVLSOKELLGPPPEAKRARGPEEEIGSPPEM-AAPASQKLS 403
Db      354  QNKPKINEDPKDAMLRFQOEIEIKKQELAAARAAG-----GGPITMPSGGSGPTQKI- 406
QY      404  PLQKLSMDP-----AMLERLLSLD-----RLLASQSGQ- 433
Db      407  -VERTEEVDPDIDAIKAQMRALAEAKKSDISTEALDKAREEAAEAAKQLQAIIDDQK 465
QY      434  -----APLSTPKRRMVLMTVEEKD-----LE-IER 460
Db      466  TEAQKKAARDALKQAEARATAGATEKEQEKAVLESRIKEMEGKIVVGVVNMLEKVD 525
QY      461  LXTKQKELEAKMLAOKAEKE 481
Db      526  LKQKSEDIKREAAIRKQBE 546

RESULT 12
H84777
probable kinesin-related cytokeleton protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: H84777
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84777
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1056 <STO>
A;Cross-references: UNIPROT:P82266; GB:AE002093; NID:g4510356; PIDN:AAD21445.1; GSPDB:G
C;Genetics:
A;Gene: At2g36200
A;Map position: 2
C;Superfamily: kinesin-related protein Eg5; kinesin motor domain homology

Query Match      22.7%; Score 561.5; DB 2; Length 1056;
Best Local Similarity 31.3%; Pred. No. 5.8e-30;
Matches 156; Conservative 86; Mismatches 152; Indels 105; Gaps 14;

QY      20  VRVAVRLRPVDGTAGASPPPCVRGMDSCSLIANWRN---HQETLKQYQDFAFYGERST 75
Db      13  VQVLLRCRPFSDDEL-RSNAPQVLTCDLQREAVSQTAGRHIDRV-FTFDKVFGPSAQ 70
QY      76  QODIYAGSVQPIRLHLEGNASVLAYGPTGAGKTHMLGSG-----PEQPGVI 123
Db      71  QKDLIDQAVVPIVNEVLEGFNCTIFAYGQTGTGKTYTMEGECRSKSAPOGGLPAEAGVI 130
QY      124  PRALMDLLOLTREBAGRPWALSVMYSYLEIYQEKVLDLLD-----ASGDLV 172
Db      131  PRAVKQIF-----DTLEGQQAQYSVKVFLELYNEEITDILLAPEDLSRVAABEKQKPLP 185
QY      173  IREDCGNLIPGLSQKPISS-----PADFERHPLPASRNRTPCATRLNQESSSHAVLLV 228
Db      186  LMEDGKGVVLVRGLEEIEIVTSANEIFTLLER-----GSSKRRTAETELNKQSSSHLSFSI 241
QY      229  KVDQERLAPFRQ-REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSFLVGLKAVDA 286
Db      242  THIKATEPEGELIKCGKLNLDLAGSENISSSGARDGAREAGEINKSLTLGRVISA 301
QY      287  LNQGLPRVYRDSKLTLLQDSLGGSAHSILIANIAPERFPYLDTVSALNPAARSKVIN 346
Db      302  LVEHLGHVYRDSKLTLLRLSDLSGRTKTCIIATVSPAVHCLSETLSTLDYAHRAKNIRN 361

```

10-10-68

[illegible]

Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D86151
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-885 <STO>
A:Cross-references: UNIPROT:Q9LPC6; GB:AE005172; NID:G8570446; PIDN:AAF76473.1; GSPDB:GN
C:Genetics:
A:Map position: 1

Query Match 22.7%; Score 560; DB 2; Length 885;
Best Local Similarity 31.5%; Pred. No. 5.7e-30;
Matches 169; Conservative 89; Mismatches 203; Indels 76; Gaps 15;
Qy 3 RCRLSKIGATRRPPARVRVAVRLRP--FVDGTAGASDPFCVR---GMDSCSLEIANRN 57
Db 44 RSSASIGADNGVPCGRVRVAVRLPRNADSVADAFDCVLPQLKELKLRKNW-- 101
Qy 58 HQETLKYQDAFYCERSTQODIYAGSVQPIRLHLLGONASVLAYGPTGAGKTHW--LG 115
Db 102 --DTETFEDEVLTAAQKRVYEVVAKPVESVLEGYNGTVMAYGQTGKTFILORLG 159
Qy 116 SPEQP--GVTPRALMDLLOLTREEGAEGRPWALSVTMSYLEIYQEKVLDLLDPASGLVI 173
Db 160 DETAARGIMVRSMDIIGTSLD-----TDSISVSYLYMETIQDLDLPTNDNIAI 212
Qy 174 REOCR--GNILPGLSQKPISSFADFERHFLPASNRNVTGATRLNQRSSRSHAVLVKY-- 230
Db 213 VEDPRTGDVSLPGATHVEIRNQONFLELLOLGETHRVAANTKLNTSSRSHAILMVHYR 272
Qy 231 -----DQERLAPFRQ-----REGKLYLIDLAGEDNRRRTGNKGLRLKESGAIN 274
Db 273 SVVNEFPVSNEMESSHFVRPSKPLVRSKLVLDLAGSERVHKSGEGHMLEAKSIN 332
Qy 275 TSLFVLGKVDALNQGLPRVPYRDSKLTLLQDLSGSAHSILIANIAPERRFYLDTVSA 334
Db 333 LSLSALGKICINATAENSHPVPLRDSKLTLLRDSFGGTARTSLIVTIGPSPRHGETTST 392
Qy 335 LNFAARSKVINRPFNTNESLQPHALGP-----VKLSQKELLG-PPEAKARGPEEE 384
Db 393 ILFGQRAKVENMLKIKEFDYKSLKLEVDKVAENRQLKAFDDDDVERINQOQN 452
Qy 385 EIGSPPEMAAPASQKLSPLQKLSMDPAMLERLLSLDR-----LLASQ 430
Db 453 RISEVEKNFAALEKEKLCQMEYMSVKLEKLSNQHNHENGKNGEVNGVVTASEF 512
Qy 431 SQGAPLLSTPKRERMVLMKTVEE-----KDLIERLKTQKELEAKMLAQKAEKE 481
Db 513 TR----LKESLENEMKLRKSAEEVSKEGEDAGITRL---QKLEDEALQKKKLEEE 562

Search completed: November 5, 2004, 18:45:35
Job time : 16.1982 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:16 ; Search time: 297.595 Seconds
(without alignments)
1813.185 Million cell updates/sec

Title: US-10-797-893-8

Perfect score: 2472

Sequence: 1 MGRCLSKIGATRRPPPARV.....LEAKLIAQAEKENHCPTM 487

Scoring table: BLOSUM62

Gap 10.0 , Gapext 0.5

Searched: 6730630 seqs, 1107998698 residues

Total number of hits satisfying chosen parameters: 6730630

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending patents AA Main:

- 1: /cgn2_6/ptodata/1/paa/US06 COMB.pcp.*
- 2: /cgn2_6/ptodata/1/paa/US06 COMB.pcp.*
- 3: /cgn2_6/ptodata/1/paa/US07 COMB.pcp.*
- 4: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 5: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 6: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 7: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 8: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 9: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 10: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 11: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 12: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 13: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 14: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 15: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 16: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 17: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 18: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 19: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 20: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 21: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 22: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 23: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 24: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 25: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 26: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 27: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 28: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 29: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 30: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 31: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 32: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 33: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 34: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 35: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 36: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2472	100.0	487	19	US-09-597-292-8
2	2472	100.0	487	33	US-10-797-893-8
3	2467	99.8	512	19	US-09-597-292-4
4	2467	99.8	512	33	US-10-797-893-4
5	2467	99.8	665	33	US-10-733-878-425
6	2309	93.4	482	27	US-10-170-2058-10738
7	2309	93.4	490	1	PCT-US03-41687-80
8	2309	93.4	490	1	PCT-US03-02234-80
9	2309	93.4	490	24	US-09-949-016-8784
10	2309	93.4	490	29	US-10-334-143-80
11	2309	93.4	490	35	US-10-940-774-8784
12	2235	90.4	559	21	US-09-724-676-90071
13	2235	90.4	559	21	US-09-724-676-90071
14	2235	90.4	588	21	US-09-724-676-90072
15	2235	90.4	588	21	US-09-724-676-90072
16	2235	90.4	597	21	US-09-724-676-90098
17	2235	90.4	597	21	US-09-724-676-90098
18	2235	90.4	626	21	US-09-724-676-90096
19	2235	90.4	626	21	US-09-724-676-90096
20	2235	90.4	627	21	US-09-724-676-90099
21	2235	90.4	627	21	US-09-724-676-90099
22	2235	90.4	656	21	US-09-724-676-90097
23	2235	90.4	656	21	US-09-724-676-90097
24	2141	86.6	460	1	PCT-US00-05988-1228
25	2141	86.6	460	24	US-09-925-300-1228
26	2128.5	86.1	627	21	US-09-724-676-90095
27	2128.5	86.1	627	21	US-09-724-676-90095
28	2081	84.2	414	27	US-10-170-2058-11425
29	2081	84.2	414	27	US-10-170-2058-11425
30	2014	81.5	665	22	US-09-791-537-149150
31	1768	71.5	346	19	US-09-597-292-6
32	1768	71.5	346	33	US-10-797-893-6
33	1763	71.3	370	19	US-09-597-292-2
34	1763	71.3	370	33	US-10-797-893-2
35	1438	58.2	651	22	US-09-791-537-149034
36	704	28.5	148	22	US-09-791-537-102346
37	687.5	27.8	147	36	US-60-182-902-537
38	673	27.2	134	36	US-60-163-233-3036
39	659	26.7	135	36	US-60-169-867-7228
40	634.5	25.7	543	30	US-10-438-246-16760
41	633	25.6	584	30	US-10-449-903-51666
42	625.5	25.3	377	27	US-10-155-881-27026
43	616	24.9	784	20	US-09-614-150A-40128
44	616	24.9	784	20	US-09-614-150A-40128
45	616	24.9	784	22	US-09-791-537-74691

ALIGNMENTS

RESULT 1
US-09-597-292-8
; Sequence 8, Application US/09597292
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09/597,292
; CURRENT FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/295,612
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Human
US-09-597-292-8

Query Match 100.0%; Score 2472; DB 19; Length 487;
 Best Local Similarity 100.0%; Pred. No. 1.3e-204; Indels 0; Gaps 0;
 Matches 487; Conservative 0; Mismatches 0;

QY 1 MGRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
 DB 1 MGRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60

QY 61 TLKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPEQ 120
 DB 61 TLKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPEQ 120

QY 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRG 180
 DB 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRG 180

QY 181 ILIPGLSOKPISSFADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 240
 DB 181 ILIPGLSOKPISSFADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 240

QY 241 QREGKLYLIDLAGSEDNRRRTGNKGLRKESGAINTSLFVLGKVDALNOGLPRVPYRDSK 300
 DB 241 QREGKLYLIDLAGSEDNRRRTGNKGLRKESGAINTSLFVLGKVDALNOGLPRVPYRDSK 300

QY 301 LTRLLQDSLGSSAHSILIANIAPERFYLDTVSALNFAARKEVINRPTNESLOPHALG 360
 DB 301 LTRLLQDSLGSSAHSILIANIAPERFYLDTVSALNFAARKEVINRPTNESLOPHALG 360

QY 361 PVKLSOKELGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420
 DB 361 PVKLSOKELGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420

QY 421 SLDRLASQSGCAPLLSTPKRERMVLMKTVEEKDLEIERLTKQKELEAKMLAKAEK 480
 DB 421 SLDRLASQSGCAPLLSTPKRERMVLMKTVEEKDLEIERLTKQKELEAKMLAKAEK 480

QY 481 ENHCPTM 487
 DB 481 ENHCPTM 487

RESULT 2

US-10-797-893-8
 ; Sequence 8, Application US/10797893
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; TITLE OF INVENTION: Novel motor proteins and methods for
 ; TITLE OF INVENTION: their use
 ; FILE REFERENCE: 1044
 ; CURRENT APPLICATION NUMBER: US/10797,893
 ; PRIOR FILING DATE: 2004-03-09
 ; PRIOR APPLICATION NUMBER: US/09/724,224
 ; PRIOR FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: 09/597,292
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 487
 ; TYPE: PRT
 ; ORGANISM: Human
 US-10-797-893-8

Query Match 100.0%; Score 2472; DB 33; Length 487;
 Best Local Similarity 100.0%; Pred. No. 1.3e-204; Indels 0; Gaps 0;
 Matches 487; Conservative 0; Mismatches 0;

QY 1 MGRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
 DB 1 MGRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60

QY 61 TLKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPEQ 120

DB 61 TLKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPEQ 120
 QY 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRG 180
 DB 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRG 180

QY 181 ILIPGLSOKPISSFADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 240
 DB 181 ILIPGLSOKPISSFADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 240

QY 241 QREGKLYLIDLAGSEDNRRRTGNKGLRKESGAINTSLFVLGKVDALNOGLPRVPYRDSK 300
 DB 241 QREGKLYLIDLAGSEDNRRRTGNKGLRKESGAINTSLFVLGKVDALNOGLPRVPYRDSK 300

QY 301 LTRLLQDSLGSSAHSILIANIAPERFYLDTVSALNFAARKEVINRPTNESLOPHALG 360
 DB 301 LTRLLQDSLGSSAHSILIANIAPERFYLDTVSALNFAARKEVINRPTNESLOPHALG 360

QY 361 PVKLSOKELGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420
 DB 361 PVKLSOKELGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420

QY 421 SLDRLASQSGCAPLLSTPKRERMVLMKTVEEKDLEIERLTKQKELEAKMLAKAEK 480
 DB 421 SLDRLASQSGCAPLLSTPKRERMVLMKTVEEKDLEIERLTKQKELEAKMLAKAEK 480

QY 481 ENHCPTM 487
 DB 481 ENHCPTM 487

RESULT 3

US-09-597-292-4
 ; Sequence 4, Application US/09597292
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; TITLE OF INVENTION: Novel motor proteins and methods for
 ; TITLE OF INVENTION: their use
 ; FILE REFERENCE: 1044
 ; CURRENT APPLICATION NUMBER: US/09/597,292
 ; PRIOR FILING DATE: 2000-06-20
 ; PRIOR APPLICATION NUMBER: US 09/295,612
 ; PRIOR FILING DATE: 1999-04-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 512
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-597-292-4

Query Match 99.8%; Score 2467; DB 19; Length 512;
 Best Local Similarity 100.0%; Pred. No. 3.9e-204; Indels 0; Gaps 0;
 Matches 486; Conservative 0; Mismatches 0;

QY 2 GRCRLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 61
 DB 27 GRCRLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 86

QY 62 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPEQ 121
 DB 87 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPEQ 146

QY 122 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRG 181
 DB 147 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRG 206

QY 182 LIPGLSOKPISSFADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 241
 DB 207 LIPGLSOKPISSFADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 266

QY 242 REGKLYLIDLAGSEDNRRRTGNKGLRKESGAINTSLFVLGKVDALNOGLPRVPYRDSK 301

Db 267 REGKYLIDLAGSDNRRTGNKGLRKSAGINTSLFVLGKVVLDALNOGLPRVPYRDSKL 326
 QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDITVSALNFAARSKEVINRPTNESLOPHALGP 361
 Db 327 TRLLQDSLGSAHSILIANIAPERRFYLDITVSALNFAARSKEVINRPTNESLOPHALGP 386
 QY 362 VKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 421
 Db 387 VKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 446
 QY 422 LDRLLASQSGCAPLLSTPKRERVMVLMKTVEEKDLEIRLTKKKELEAKMLAQAEKE 481
 Db 447 LDRLLASQSGCAPLLSTPKRERVMVLMKTVEEKDLEIRLTKKKELEAKMLAQAEKE 506
 QY 482 NHCPTM 487
 Db 507 NHCPTM 512

RESULT 4
 US-10-797-893-4
 ; Sequence 4, Application US/10797893
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; TITLE OF INVENTION: Novel motor proteins and methods for
 ; FILE REFERENCE: 1044
 ; CURRENT APPLICATION NUMBER: US/10/797,893
 ; PRIOR FILING DATE: 2004-03-09
 ; PRIOR FILING DATE: 2000-11-28
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 512
 ; TYPE: PRT
 ; ORGANISM: Human
 US-10-797-893-4

Query Match 99.8%; Score 2467; DB 33; Length 512;
 Best Local Similarity 100.0%; Pred. No. 3.9e-204;
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEITANRNHOET 61
 Db 27 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEITANRNHOET 86
 QY 62 LKQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 121
 Db 87 LKQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 146
 QY 122 VIPRALMDLLQLTREAGRPWALSVMYSLEYIQEKVLDLDDPPASGDLVIREDCRGN 181
 Db 147 VIPRALMDLLQLTREAGRPWALSVMYSLEYIQEKVLDLDDPPASGDLVIREDCRGN 206
 QY 182 LIPGLSOKPTSSFADFERHFLPASRNTVGATRLNQRSSRSRSHAVLLVKVDORERLAPFRQ 241
 Db 207 LIPGLSOKPTSSFADFERHFLPASRNTVGATRLNQRSSRSRSHAVLLVKVDORERLAPFRQ 266
 QY 242 REGKYLIDLAGSDNRRTGNKGLRKSAGINTSLFVLGKVVLDALNOGLPRVPYRDSKL 301
 Db 267 REGKYLIDLAGSDNRRTGNKGLRKSAGINTSLFVLGKVVLDALNOGLPRVPYRDSKL 326
 QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDITVSALNFAARSKEVINRPTNESLOPHALGP 361
 Db 327 TRLLQDSLGSAHSILIANIAPERRFYLDITVSALNFAARSKEVINRPTNESLOPHALGP 386
 QY 362 VKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 421
 Db 387 VKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 446

QY 422 LDRLLASQSGCAPLLSTPKRERVMVLMKTVEEKDLEIRLTKKKELEAKMLAQAEKE 481
 Db 447 LDRLLASQSGCAPLLSTPKRERVMVLMKTVEEKDLEIRLTKKKELEAKMLAQAEKE 506
 QY 482 NHCPTM 487
 Db 507 NHCPTM 512
 RESULT 5
 US-10-733-878-425
 ; Sequence 425, Application US/10733878
 ; GENERAL INFORMATION:
 ; APPLICANT: Jean-Philippe Girard
 ; APPLICANT: Francois Amalric
 ; APPLICANT: Myriam Roussigne
 ; APPLICANT: Thomas Clouaire
 ; TITLE OF INVENTION: THAP PROTEINS AS NUCLEAR RECEPTORS FOR
 ; TITLE OF INVENTION: CHEMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
 ; TITLE OF INVENTION: PROLIFERATION AND CELL DIFFERENTIATION
 ; FILE REFERENCE: BIOBANK 012A
 ; CURRENT APPLICATION NUMBER: US/10/733,878
 ; CURRENT FILING DATE: 2003-12-10
 ; PRIOR APPLICATION NUMBER: 60/432699
 ; PRIOR FILING DATE: 2002-12-10
 ; PRIOR APPLICATION NUMBER: 60/485027
 ; PRIOR FILING DATE: 2003-07-03
 ; NUMBER OF SEQ ID NOS: 535
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 425
 ; LENGTH: 665
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-733-878-425

Query Match 99.8%; Score 2467; DB 33; Length 665;
 Best Local Similarity 100.0%; Pred. No. 5.9e-204;
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEITANRNHOET 61
 Db 26 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEITANRNHOET 85
 QY 62 LKQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 121
 Db 86 LKQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 145
 QY 122 VIPRALMDLLQLTREAGRPWALSVMYSLEYIQEKVLDLDDPPASGDLVIREDCRGN 181
 Db 146 VIPRALMDLLQLTREAGRPWALSVMYSLEYIQEKVLDLDDPPASGDLVIREDCRGN 205
 QY 182 LIPGLSOKPTSSFADFERHFLPASRNTVGATRLNQRSSRSRSHAVLLVKVDORERLAPFRQ 241
 Db 206 LIPGLSOKPTSSFADFERHFLPASRNTVGATRLNQRSSRSRSHAVLLVKVDORERLAPFRQ 265
 QY 242 REGKYLIDLAGSDNRRTGNKGLRKSAGINTSLFVLGKVVLDALNOGLPRVPYRDSKL 301
 Db 266 REGKYLIDLAGSDNRRTGNKGLRKSAGINTSLFVLGKVVLDALNOGLPRVPYRDSKL 325
 QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDITVSALNFAARSKEVINRPTNESLOPHALGP 361
 Db 326 TRLLQDSLGSAHSILIANIAPERRFYLDITVSALNFAARSKEVINRPTNESLOPHALGP 385
 QY 362 VKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 421
 Db 386 VKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 445
 QY 422 LDRLLASQSGCAPLLSTPKRERVMVLMKTVEEKDLEIRLTKKKELEAKMLAQAEKE 481
 Db 446 LDRLLASQSGCAPLLSTPKRERVMVLMKTVEEKDLEIRLTKKKELEAKMLAQAEKE 505
 QY 482 NHCPTM 487

Db 506 NHCPTM 511

RESULT 6
US-10-170-205E-10738
; Sequence 80, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10738
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-10738

Query Match 93.4%; Score 2309; DB 27; Length 482;
Best Local Similarity 100.0%; Pred. No. 1.8e-190;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61
Db 26 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 85
QY 62 LKQFDIFYGERSTQDDIYAGSVQPIRLHLEQONASVLAAYGPTGAGKTHMLGSPQPG 121
Db 86 LKQFDIFYGERSTQDDIYAGSVQPIRLHLEQONASVLAAYGPTGAGKTHMLGSPQPG 145
QY 122 VIPRALMDLLQLTREGEAGRPWALSVTMSYLIYQEKVLDLDPASGDLVIREDCRGN 181
Db 146 VIPRALMDLLQLTREGEAGRPWALSVTMSYLIYQEKVLDLDPASGDLVIREDCRGN 205
QY 182 LIPGLSQKPISSFADFERHFLPASRNRVTGATRLNQRSSSHAVLLVKVDORERLAPRQ 241
Db 206 LIPGLSQKPISSFADFERHFLPASRNRVTGATRLNQRSSSHAVLLVKVDORERLAPRQ 265
QY 242 REGKYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVPYRDSKL 301
Db 266 REGKYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVPYRDSKL 325
QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPTNESLOPHALGP 361
Db 326 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPTNESLOPHALGP 385
QY 362 VKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 421
Db 386 VKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 445
QY 422 LDRLLASOGSOGAPLLSTPKERMVLMKTVEEKDL 456
Db 446 LDRLLASOGSOGAPLLSTPKERMVLMKTVEEKDL 480

RESULT 7
PCT-US02-41687-80
; Sequence 80, Application PC/TUS0241687
; GENERAL INFORMATION:
; APPLICANT: SUGEN, INC.
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD
; FILE REFERENCE: 038602/1544
; CURRENT APPLICATION NUMBER: PCT/US02/41687
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: 60/343,169
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 80

Query Match 93.4%; Score 2309; DB 1; Length 490;
Best Local Similarity 100.0%; Pred. No. 1.8e-190;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61
Db 34 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 93
QY 62 LKQFDIFYGERSTQDDIYAGSVQPIRLHLEQONASVLAAYGPTGAGKTHMLGSPQPG 121
Db 94 LKQFDIFYGERSTQDDIYAGSVQPIRLHLEQONASVLAAYGPTGAGKTHMLGSPQPG 153

RESULT 8
PCT-US03-02234-80
; Sequence 80, Application PC/TUS0302234
; GENERAL INFORMATION:
; APPLICANT: GRIGORIEV, IGOR V.
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD
; FILE REFERENCE: 038602/1571
; CURRENT APPLICATION NUMBER: PCT/US03/02234
; CURRENT FILING DATE: 2003-01-28
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 80
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-02234-80

Query Match 93.4%; Score 2309; DB 1; Length 490;
Best Local Similarity 100.0%; Pred. No. 1.8e-190;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61
Db 34 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 93
QY 62 LKQFDIFYGERSTQDDIYAGSVQPIRLHLEQONASVLAAYGPTGAGKTHMLGSPQPG 121
Db 94 LKQFDIFYGERSTQDDIYAGSVQPIRLHLEQONASVLAAYGPTGAGKTHMLGSPQPG 153

```
Qy 122 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQKVDLLDLPASGDLVIREDCRNI 181
Db 154 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQKVDLLDLPASGDLVIREDCRNI 213
Qy 182 LIPGLSQKPISSFADFERRHFLPASRNTVGTATRLNQRSSSHAVLLVKVDORERLAPRQ 241
Db 214 LIPGLSQKPISSFADFERRHFLPASRNTVGTATRLNQRSSSHAVLLVKVDORERLAPRQ 273
Qy 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVDVLDLQGLPRVPRYRDSKL 301
Db 274 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVDVLDLQGLPRVPRYRDSKL 333
Qy 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVETINRPTNESLOPHALGP 361
Db 334 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVETINRPTNESLOPHALGP 393
Qy 362 VKLSQKELGPPKAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 421
Db 394 VKLSQKELGPPKAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 453
Qy 422 LDRLLASQSGAGAPLLSTPKRERVMKMTVEEKDL 456
Db 454 LDRLLASQSGAGAPLLSTPKRERVMKMTVEEKDL 488

RESULT 9
US-09-949-016-8784
; Sequence 8784, Application US/09949016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYOMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8784
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8784

Query Match 93.4%; Score 2309; DB 24; Length 490;
Best Local Similarity 100.0%; Pred. No. 1.8e-190;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61
Db 34 GRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 93
Qy 62 LKYOFAFYGERSTQDDIYAGSVQPIRLHLEGNASVLAGPTGAGKTHMLGSPRQ 121
Db 94 LKYOFAFYGERSTQDDIYAGSVQPIRLHLEGNASVLAGPTGAGKTHMLGSPRQ 153
Qy 122 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQKVDLLDLPASGDLVIREDCRNI 181
Db 154 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQKVDLLDLPASGDLVIREDCRNI 213
Qy 182 LIPGLSQKPISSFADFERRHFLPASRNTVGTATRLNQRSSSHAVLLVKVDORERLAPRQ 241
Db 214 LIPGLSQKPISSFADFERRHFLPASRNTVGTATRLNQRSSSHAVLLVKVDORERLAPRQ 273
Qy 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVDVLDLQGLPRVPRYRDSKL 301
Db 274 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVDVLDLQGLPRVPRYRDSKL 333
Qy 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVETINRPTNESLOPHALGP 361
Db 334 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVETINRPTNESLOPHALGP 393
Qy 362 VKLSQKELGPPKAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 421
Db 394 VKLSQKELGPPKAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 453
Qy 422 LDRLLASQSGAGAPLLSTPKRERVMKMTVEEKDL 456
Db 454 LDRLLASQSGAGAPLLSTPKRERVMKMTVEEKDL 488

RESULT 11
US-10-940-774-8784
```

```
Qy 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVETINRPTNESLOPHALGP 361
Db 334 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVETINRPTNESLOPHALGP 393
Qy 362 VKLSQKELGPPKAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 421
Db 394 VKLSQKELGPPKAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 453
Qy 422 LDRLLASQSGAGAPLLSTPKRERVMKMTVEEKDL 456
Db 454 LDRLLASQSGAGAPLLSTPKRERVMKMTVEEKDL 488

RESULT 10
US-10-334-143-80
; Sequence 80, Application US/10334143
; GENERAL INFORMATION:
; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
; FILE REFERENCE: 038602/1543
; CURRENT APPLICATION NUMBER: US/10/334,143
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: 60/343,169
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 80
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-334-143-80

Query Match 93.4%; Score 2309; DB 29; Length 490;
Best Local Similarity 100.0%; Pred. No. 1.8e-190;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61
Db 34 GRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 93
Qy 62 LKYOFAFYGERSTQDDIYAGSVQPIRLHLEGNASVLAGPTGAGKTHMLGSPRQ 121
Db 94 LKYOFAFYGERSTQDDIYAGSVQPIRLHLEGNASVLAGPTGAGKTHMLGSPRQ 153
Qy 122 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQKVDLLDLPASGDLVIREDCRNI 181
Db 154 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQKVDLLDLPASGDLVIREDCRNI 213
Qy 182 LIPGLSQKPISSFADFERRHFLPASRNTVGTATRLNQRSSSHAVLLVKVDORERLAPRQ 241
Db 214 LIPGLSQKPISSFADFERRHFLPASRNTVGTATRLNQRSSSHAVLLVKVDORERLAPRQ 273
Qy 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVDVLDLQGLPRVPRYRDSKL 301
Db 274 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVDVLDLQGLPRVPRYRDSKL 333
Qy 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVETINRPTNESLOPHALGP 361
Db 334 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVETINRPTNESLOPHALGP 393
Qy 362 VKLSQKELGPPKAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 421
Db 394 VKLSQKELGPPKAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 453
Qy 422 LDRLLASQSGAGAPLLSTPKRERVMKMTVEEKDL 456
Db 454 LDRLLASQSGAGAPLLSTPKRERVMKMTVEEKDL 488

RESULT 11
US-10-940-774-8784
```

Fri Nov 12 12:26:28 2004

```

; ORGANISM: Homo sapiens
; US-09-724-676-90071

Query Match      90.4%; Score 2235; DB 21; Length 559;
Best Local Similarity 99.8%; Pred. No. 5.9e-184;
Matches 442; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 45 MDSCSLEIANWRNHQETLKYQFDAPFYGERSTQODIYAGSVQPIRLHLLGQNASVLYAYGP 104
Db 1 MDSCSLEIANWRNHQETLKYQFDAPFYGERSTQODIYAGSVQPIRLHLLGQNASVLYAYGP 60

Qy 105 TGAGKTHTMLGSPQPGVIPRALMDLLQLTREBGAERPWALSVTMSYLEIYQEKVLDLL 164
Db 61 TGAGKTHTMLGSPQPGVIPRALMDLLQLTREBGAERPWALSVTMSYLEIYQEKVLDLL 120

Qy 165 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNTVGTATLNRSSRSHA 224
Db 121 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNTVGTATLNRSSRSHA 180

Qy 225 VLLVKVDORERLAPPROEGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVV 284
Db 181 VLLVKVDORERLAPPROEGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVV 240

Qy 285 DALNOGLPRVPRYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEY 344
Db 241 DALNOGLPRVPRYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEY 300

Qy 345 INRPFTNESLOPHALGPVKLSQKELLGPPAKRARGPEEIEIGSPPEMAAPASASQKLSP 404
Db 301 INRPFTNESLOPHALGPVKLSQKELLGPPAKRARGPEEIEIGSPPEMAAPASASQKLSP 360

Qy 405 LQKSSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDLEIERLTK 464
Db 361 LQKSSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDLEIERLTK 420

Qy 465 QKELEAKMLAQKAEKENHCPTM 487
Db 421 QKELEAKMLAQKAEKENHCPTM 443

RESULT 13
US-09-724-676A-90071
; Sequence 90071, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 90071
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-724-676A-90071

Query Match      90.4%; Score 2235; DB 21; Length 559;
Best Local Similarity 99.8%; Pred. No. 5.9e-184;
Matches 442; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 45 MDSCSLEIANWRNHQETLKYQFDAPFYGERSTQODIYAGSVQPIRLHLLGQNASVLYAYGP 104
Db 1 MDSCSLEIANWRNHQETLKYQFDAPFYGERSTQODIYAGSVQPIRLHLLGQNASVLYAYGP 60

Qy 105 TGAGKTHTMLGSPQPGVIPRALMDLLQLTREBGAERPWALSVTMSYLEIYQEKVLDLL 164
Db 61 TGAGKTHTMLGSPQPGVIPRALMDLLQLTREBGAERPWALSVTMSYLEIYQEKVLDLL 120

Qy 165 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNTVGTATLNRSSRSHA 224
Db 121 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNTVGTATLNRSSRSHA 180

; Sequence 8784, Application US/10940774
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/10/940,774
; CURRENT FILING DATE: 2004-09-15
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 8784
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Human
; US-10-940-774-8784

Query Match      93.4%; Score 2309; DB 35; Length 490;
Best Local Similarity 100.0%; Pred. No. 1.8e-190;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRCLSKIGATRRPPPARVRVAVLRPFVDGTAGSDPPCVRGMDSCSLEIANWRNHQET 61
Db 34 GRCLSKIGATRRPPPARVRVAVLRPFVDGTAGSDPPCVRGMDSCSLEIANWRNHQET 93

Qy 62 LKYQFDAPFYGERSTQODIYAGSVQPIRLHLLGQNASVLYAYGPTGAGKTHTMLGSPQPG 121
Db 94 LKYQFDAPFYGERSTQODIYAGSVQPIRLHLLGQNASVLYAYGPTGAGKTHTMLGSPQPG 153

Qy 122 VIPRALMDLLQLTREBGAERPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN 181
Db 154 VIPRALMDLLQLTREBGAERPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN 213

Qy 182 LIPGLSQKPISSPADFERHFLPASRNTVGTATLNRSSRSHAVLLVKVDORERLAPPRO 241
Db 214 LIPGLSQKPISSPADFERHFLPASRNTVGTATLNRSSRSHAVLLVKVDORERLAPPRO 273

Qy 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVVLDALNOGLPRVPRYRDSKL 301
Db 274 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVVLDALNOGLPRVPRYRDSKL 333

Qy 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEYINRPFTNESLOPHALGP 361
Db 334 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEYINRPFTNESLOPHALGP 393

Qy 362 VKLSQKELLGPPAKRARGPEEIEIGSPPEMAAPASASQKLSPLOKSSMDPAMLERLLS 421
Db 394 VKLSQKELLGPPAKRARGPEEIEIGSPPEMAAPASASQKLSPLOKSSMDPAMLERLLS 453

Qy 422 LDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDL 456
Db 454 LDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDL 488

RESULT 12
US-09-724-676-90071
; Sequence 90071, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 90071
; LENGTH: 559
; TYPE: PRT

```

```
Qy 225 VLLVKVDQERLAPPRQEGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKV 284
Db 181 VLLVKVDQERLAPPRQEGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKV 240
Qy 285 DALNOGLPRVPRDSKLTLLQDSLGSAHSILITANIAPERRFYLDTVSALNFAARSKEV 344
Db 241 DALNOGLPRVPRDSKLTLLQDSLGSAHSILITANIAPERRFYLDTVSALNFAARSKEV 300
Qy 345 INRPFTNESLQPHALGPVKLSQKELLGPPPEAKRARGPEEBEIGSPPEMAAPASASOKLSP 404
Db 301 INRPFTNESLQPHALGPVKLSQKELLGPPPEAKRARGPEEBEIGSPPEMAAPASASOKLSP 360
Qy 405 LQKLSSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDLEIERLTK 464
Db 361 LQKLSSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDLEIERLTK 420
Qy 465 QKELEAKMLAQKAEKENHCPTM 487
Db 421 QKELEAKMLAQKAEKENHCPTM 443

RESULT 14
US-09-724-676-90072
; Sequence 90072, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 90072
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-90072

Query Match 90.4%; Score 2235; DB 21; Length 588;
Best Local Similarity 99.8%; Pred. No. 6.4e-184;
Matches 442; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 45 MDSCSLEIANWRNHQETLKQFDIFYGERSTQQDIYAGSVQPIRLHLLLEGQNASVLYGYP 104
Db 1 MDSCSLEIANWRNHQETLKQFDIFYGERSTQQDIYAGSVQPIRLHLLLEGQNASVLYGYP 60
Qy 105 TGAGKTHMLGSPQPGVPIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDLL 164
Db 61 TGAGKTHMLGSPQPGVPIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDLL 120
Qy 165 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNRRTVGATRLNQRSSSHA 224
Db 121 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNRRTVGATRLNQRSSSHA 180
Qy 225 VLLVKVDQERLAPPRQEGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKV 284
Db 181 VLLVKVDQERLAPPRQEGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKV 240
Qy 285 DALNOGLPRVPRDSKLTLLQDSLGSAHSILITANIAPERRFYLDTVSALNFAARSKEV 344
Db 241 DALNOGLPRVPRDSKLTLLQDSLGSAHSILITANIAPERRFYLDTVSALNFAARSKEV 300
Qy 345 INRPFTNESLQPHALGPVKLSQKELLGPPPEAKRARGPEEBEIGSPPEMAAPASASOKLSP 404
Db 301 INRPFTNESLQPHALGPVKLSQKELLGPPPEAKRARGPEEBEIGSPPEMAAPASASOKLSP 360
Qy 405 LQKLSSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDLEIERLTK 464
Db 361 LQKLSSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDLEIERLTK 420
Qy 465 QKELEAKMLAQKAEKENHCPTM 487
Db 421 QKELEAKMLAQKAEKENHCPTM 443

RESULT 14
US-09-724-676-90072
; Sequence 90072, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 90072
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-90072

Query Match 90.4%; Score 2235; DB 21; Length 588;
Best Local Similarity 99.8%; Pred. No. 6.4e-184;
Matches 442; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 45 MDSCSLEIANWRNHQETLKQFDIFYGERSTQQDIYAGSVQPIRLHLLLEGQNASVLYGYP 104
Db 1 MDSCSLEIANWRNHQETLKQFDIFYGERSTQQDIYAGSVQPIRLHLLLEGQNASVLYGYP 60
Qy 105 TGAGKTHMLGSPQPGVPIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDLL 164
Db 61 TGAGKTHMLGSPQPGVPIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDLL 120
Qy 165 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNRRTVGATRLNQRSSSHA 224
Db 121 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNRRTVGATRLNQRSSSHA 180
Qy 225 VLLVKVDQERLAPPRQEGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKV 284
Db 181 VLLVKVDQERLAPPRQEGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKV 240
Qy 285 DALNOGLPRVPRDSKLTLLQDSLGSAHSILITANIAPERRFYLDTVSALNFAARSKEV 344
Db 241 DALNOGLPRVPRDSKLTLLQDSLGSAHSILITANIAPERRFYLDTVSALNFAARSKEV 300
Qy 345 INRPFTNESLQPHALGPVKLSQKELLGPPPEAKRARGPEEBEIGSPPEMAAPASASOKLSP 404
Db 301 INRPFTNESLQPHALGPVKLSQKELLGPPPEAKRARGPEEBEIGSPPEMAAPASASOKLSP 360
Qy 405 LQKLSSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDLEIERLTK 464
Db 361 LQKLSSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDLEIERLTK 420
Qy 465 QKELEAKMLAQKAEKENHCPTM 487
Db 421 QKELEAKMLAQKAEKENHCPTM 443
```

RESULT 15

```
US-09-724-676A-90072
; Sequence 90072, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 90072
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-90072
```

```
Query Match 90.4%; Score 2235; DB 21; Length 588;
Best Local Similarity 99.8%; Pred. No. 6.4e-184;
Matches 442; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 45 MDSCSLEIANWRNHQETLKQFDIFYGERSTQQDIYAGSVQPIRLHLLLEGQNASVLYGYP 104
Db 1 MDSCSLEIANWRNHQETLKQFDIFYGERSTQQDIYAGSVQPIRLHLLLEGQNASVLYGYP 60
Qy 105 TGAGKTHMLGSPQPGVPIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDLL 164
Db 61 TGAGKTHMLGSPQPGVPIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDLL 120
Qy 165 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNRRTVGATRLNQRSSSHA 224
Db 121 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNRRTVGATRLNQRSSSHA 180
Qy 225 VLLVKVDQERLAPPRQEGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKV 284
Db 181 VLLVKVDQERLAPPRQEGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKV 240
Qy 285 DALNOGLPRVPRDSKLTLLQDSLGSAHSILITANIAPERRFYLDTVSALNFAARSKEV 344
Db 241 DALNOGLPRVPRDSKLTLLQDSLGSAHSILITANIAPERRFYLDTVSALNFAARSKEV 300
Qy 345 INRPFTNESLQPHALGPVKLSQKELLGPPPEAKRARGPEEBEIGSPPEMAAPASASOKLSP 404
Db 301 INRPFTNESLQPHALGPVKLSQKELLGPPPEAKRARGPEEBEIGSPPEMAAPASASOKLSP 360
Qy 405 LQKLSSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDLEIERLTK 464
Db 361 LQKLSSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDLEIERLTK 420
Qy 465 QKELEAKMLAQKAEKENHCPTM 487
Db 421 QKELEAKMLAQKAEKENHCPTM 443
```

Search completed: November 5, 2004, 19:04:21
Job time : 299.595 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 ; Search time 99.9557 Seconds
(without alignments)
1747.786 Million cell updates/sec

Title: US-10-797-893-8

Perfect score: 2472

Sequence: 1 MGRCLSKIGATRRPPARV.....LEAKMLQAQAEKENHCPTM 487

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2472	100.0	487	7	ADC23344 Human kin
2	2472	100.0	487	8	ADQ60234 Human mic
3	2467	99.8	512	7	ADC23340 Human kin
4	2467	99.8	512	8	ADQ60230 Human mic
5	2467	99.8	665	8	ADQ09240 Human KNS
6	2309	93.4	490	7	ADK40973 Novel hum
7	2141	86.6	460	3	AAB56650 Human pro
8	1768	71.5	346	7	ADC23342 Human kin
9	1768	71.5	346	8	ADQ60232 Human mic
10	1763	71.3	370	7	ADC23338 Human kin
11	1763	71.3	370	8	ADQ60228 Human mic
12	616	24.9	784	4	ABB71112 Drosophil
13	593	24.0	677	4	ABB65183 Drosophil
14	578.5	23.4	1237	7	ADK40926 Novel hum
15	574.5	23.2	1232	7	ADK49938 Human lun
16	574.5	23.2	1232	7	ADK49938 Human lun
17	574.5	23.2	1232	7	ADJ18924 Human dis
18	574.5	23.2	1232	7	ADJ70347 Human hea
19	573.5	23.2	1234	5	ADK40901 Novel hum
20	572	23.1	522	5	ABG70993 Human HsK
21	568.5	23.0	473	5	ABG70992 Human tar
22	568	23.0	898	4	ABU53123 Intracell
23	568	23.0	898	5	AAU76957 Novel hum
24	568	23.0	898	7	ADC31082 Human nov
25	568	23.0	898	7	AD115915 Human PP

ALIGNMENTS

RESULT 1

ADC23344
ID ADC23344 standard; protein; 487 AA.

XX AC ADC23344;

XX DT 18-DEC-2003 (first entry)

XX DE Human kinesin-like DNA binding protein (KID) (SeqID 8).

XX KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID; cytosolic; cardiac; immunomodulator; antiinflammatory; gene therapy; cancer; hyperplasia; restenosis; cellular proliferation disorder; cardiac hypertrophy; immune disorder; inflammation.

XX OS Homo sapiens.

XX PN US6387644-B1.

XX PD 14-MAY-2002.

XX PF 28-NOV-2000; 2000US-00724224.

XX PR 20-APR-1999; 99US-00295612.

XX PR 20-JUN-2000; 2000US-00597292.

XX PA (CYTO-) CYTOKINETICS INC.

XX PI Beraud C;

XX WPI; 2003-706919/67.

XX N-PSDB; ADC23343.

XX PT Identifying a candidate agent as modulator of function of a target protein for treating cellular proliferation disorders by adding a candidate agent to a mixture of the target protein that directly/indirectly produces ADP or phosphate.

XX PS Claim 1; SEQ ID NO 8; 26pp; English.

XX CC This invention relates to a novel method for high throughput screening systems used to identify compounds for the treatment of cellular proliferation disorders. Specifically, it refers to candidate agents that are capable of modulating the activity of target proteins having motor domains, such that the target protein directly or indirectly produces ADP or phosphate. Furthermore, this activity can be determined using fluorescence or absorbance readouts. The present invention describes a method that identifies modulators of the target protein, which is a

CC kinesin-like DNA binding protein (known as KID) as cytoskeletal, cardinals,
CC immunomodulators and antiinflammatory. Accordingly, through gene
CC therapy, they can be used for the treatment of cancer, hyperplasias,
CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This
CC polypeptide sequence is human KID protein (SeqID 8) of the invention.
XX
SQ Sequence 487 AA;

Query Match 100.0%; Score 2472; DB 7; Length 487;
Best Local Similarity 100.0%; Pred. No. 6e-224;
Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
DB 1 MGRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
QY 61 TLKYQDFAFYGERSTQODIYAGSVQPIRLHLEGQNASVLYAGTCAKTHMLGSPROP 120
DB 61 TLKYQDFAFYGERSTQODIYAGSVQPIRLHLEGQNASVLYAGTCAKTHMLGSPROP 120
QY 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRN 180
DB 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRN 180
QY 181 ILIPGLSQPISSFADFERHFLPASNRRTVGATRLNQSRSHAVLLVKVDQERLAPPR 240
DB 181 ILIPGLSQPISSFADFERHFLPASNRRTVGATRLNQSRSHAVLLVKVDQERLAPPR 240
QY 241 QREGKLYLIDLAGSENRRTGNKGLRKESGAINTSFLVGLKVVDALNQLPRVPYRDSK 300
DB 241 QREGKLYLIDLAGSENRRTGNKGLRKESGAINTSFLVGLKVVDALNQLPRVPYRDSK 300
QY 301 LTRLLQDSLGGSAHSILIANIAPERRFYLDTSALNFAARKEVINRPTNESLOPHALG 360
DB 301 LTRLLQDSLGGSAHSILIANIAPERRFYLDTSALNFAARKEVINRPTNESLOPHALG 360
QY 361 PVKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420
DB 361 PVKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420
QY 421 SLDRLLASQSGQAPLLSTPKRRRWLMKTVEEKDLEIERLTKQKELEAKMLAQKAEK 480
DB 421 SLDRLLASQSGQAPLLSTPKRRRWLMKTVEEKDLEIERLTKQKELEAKMLAQKAEK 480
QY 481 ENHCPTM 487
DB 481 ENHCPTM 487

RESULT 2
ADQ60234
ID ADQ60234 standard; protein; 487 AA.
XX AC ADQ60234;
XX AC ADQ60234;
DT 23-SEP-2004 (first entry)
DE Human microtubule motor protein #4.

XX Human; microtubule motor protein; cellular proliferation disorder;
KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;
KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;
KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.
XX Homo sapiens.
XX US6762043-B1.
XX 13-JUL-2004.
XX 06-MAR-2002; 2002US-00093317.
XX 20-APR-1999; 99US-00295612.
PR

PR 20-JUN-2000; 2000US-00597292.
PR 28-NOV-2000; 2000US-00724224.
XX
PA (CYTO-) CYTOKINETICS INC.
XX
PI Beraud C;
XX
DR WPI; 2004-532491/51.
XX N-ESDB; ADQ60233.

New isolated microtubule motor protein, useful for screening modulators
for treating cellular proliferation disorders such as cancer,
hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
inflammation.

Claim 1; SEQ ID NO 8; 26pp; English.

The invention relates to human microtubule motor proteins and the nucleic
acids encoding them. The invention also relates to a method of screening
for modulators of a motor protein which has microtubule stimulated ATPase
activity, a method of testing for ATPase activity of microtubule motor
protein or act as a modulator of the binding characteristics of
biological activity of a target protein, modulators of the target
protein, and methods of treating cellular proliferation disorders such as
cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
and inflammation, for treating disorders associated with kinesin-like DNA
binding protein (KID) and for inhibiting KID. The sequences are used for
screening for modulators of motor proteins useful for treating cellular
proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
hypertrophy, immune disorders and inflammation, for treating disorders
associated with KID and for inhibiting KID and for treating autoimmune
diseases, arthritis, graft rejection, inflammatory bowel disease and
proliferation induced after medical procedures including surgery and
angioplasty. This sequence represents a human microtubule motor protein
of the invention.

Sequence 487 AA;

Query Match 100.0%; Score 2472; DB 8; Length 487;
Best Local Similarity 100.0%; Pred. No. 6e-224;
Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
DB 1 MGRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
QY 61 TLKYQDFAFYGERSTQODIYAGSVQPIRLHLEGQNASVLYAGTCAKTHMLGSPROP 120
DB 61 TLKYQDFAFYGERSTQODIYAGSVQPIRLHLEGQNASVLYAGTCAKTHMLGSPROP 120
QY 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRN 180
DB 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRN 180
QY 181 ILIPGLSQPISSFADFERHFLPASNRRTVGATRLNQSRSHAVLLVKVDQERLAPPR 240
DB 181 ILIPGLSQPISSFADFERHFLPASNRRTVGATRLNQSRSHAVLLVKVDQERLAPPR 240
QY 241 QREGKLYLIDLAGSENRRTGNKGLRKESGAINTSFLVGLKVVDALNQLPRVPYRDSK 300
DB 241 QREGKLYLIDLAGSENRRTGNKGLRKESGAINTSFLVGLKVVDALNQLPRVPYRDSK 300
QY 301 LTRLLQDSLGGSAHSILIANIAPERRFYLDTSALNFAARKEVINRPTNESLOPHALG 360
DB 301 LTRLLQDSLGGSAHSILIANIAPERRFYLDTSALNFAARKEVINRPTNESLOPHALG 360
QY 361 PVKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420
DB 361 PVKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420
QY 421 SLDRLLASQSGQAPLLSTPKRRRWLMKTVEEKDLEIERLTKQKELEAKMLAQKAEK 480

Db 421 SLDRLLASQSGGAPLLSTPKRERVMKMTVEEKDLEIERLTKQKELEAKQLAKAEK 480

Qy 481 ENHCPTM 487

Db 481 ENHCPTM 487

RESULT 3

ADC23340

ID ADC23340 standard; protein; 512 AA.

XX

AC ADC23340;

XX

DT 18-DEC-2003 (first entry)

XX

DE Human kinesin-like DNA binding protein (KID) (SeqID 4).

XX

XX human; enzyme; motor domain; kinesin-like DNA binding protein; KID;

KW cytosolic; cardiac; immunomodulator; antiinflammatory; gene therapy;

KW cancer; hyperplasia; restenosis; cellular proliferation disorder;

KW cardiac hypertrophy; immune disorder; inflammation.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Misc-difference 2 /note= "Encoded by CA"

FT

XX US6387644-B1.

XX

PD 14-MAY-2002.

XX

XX 28-NOV-2000; 2000US-00724224.

XX

XX 20-APR-1999; 99US-00295612.

PR

XX 20-JUN-2000; 2000US-00597292.

XX

PA (CYTO-) CYTOKINETICS INC.

XX

PI Beraud C;

XX

DR WPI; 2003-706919/67.

DR

DR N-PSDB; ADC23339.

XX

XX Identifying a candidate agent as modulator of function of a target

PT protein for treating cellular proliferation disorders by adding a

PT candidate agent to a mixture of the target protein that

PT directly/indirectly produces ADP or phosphate.

XX

PS Claim 1; SEQ ID NO 4; 26pp; English.

XX

CC This invention relates to a novel method for high throughput screening

CC systems used to identify compounds for the treatment of cellular

CC proliferation disorders. Specifically, it refers to candidate agents that

CC are capable of modulating the activity of target proteins having motor

CC domains, such that the target protein directly or indirectly produces ADP

CC or phosphate. Furthermore, this activity can be determined using

CC fluorescence or absorbance readouts. The present invention describes a

CC method that identifies modulators of the target protein, which is a

CC kinesin-like DNA binding protein (known as KID) as cytosolic, cardiac,

CC immunomodulators and antiinflammatory. Accordingly, through gene

CC therapy, they can be used for the treatment of cancer, hyperplasias,

CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This

CC polypeptide sequence is human KID protein (SeqID 4) of the invention.

XX

SQ Sequence 512 AA;

Query Match 99.8%; Score 2467; DB 7; Length 512;

Best Local Similarity 100.0%; Pred. No. 1.9e-223;

Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRCLSKIGATRRPPPARVAVRLRPFVDTAGASDPPCCVRGMDSCSLEIANWRNHQET 61

Db 27 GRCLSKIGATRRPPPARVAVRLRPFVDTAGASDPPCCVRGMDSCSLEIANWRNHQET 86

Qy 62 LKYOFAFYGERSTQDDIYAGSVQPIILRHLLLEGQNASVLYAGPTGAGKTHMTLGSPEQPG 121

Db 87 LKYOFAFYGERSTQDDIYAGSVQPIILRHLLLEGQNASVLYAGPTGAGKTHMTLGSPEQPG 146

Qy 122 VIPRALMDLLQLTREEGAERPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN 181

Db 147 VIPRALMDLLQLTREEGAERPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN 206

Qy 182 LIPGLSOKPISSFADFERHFLPASRNRVTGATRLNQRSSRSHAVLLVKVDQERLAPPRQ 241

Db 207 LIPGLSOKPISSFADFERHFLPASRNRVTGATRLNQRSSRSHAVLLVKVDQERLAPPRQ 266

Qy 242 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSFLVGLKVVDALNOGLPRVPYRSKL 301

Db 267 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSFLVGLKVVDALNOGLPRVPYRSKL 326

Qy 302 TRLLQDSLGGSAHSILIANIAPERFYLDVTVSALNFAARKEVINRFTNESLQPHALGP 361

Db 327 TRLLQDSLGGSAHSILIANIAPERFYLDVTVSALNFAARKEVINRFTNESLQPHALGP 386

Qy 362 VKLSQKELLGPPPEAKRARGPEEBEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 421

Db 387 VKLSQKELLGPPPEAKRARGPEEBEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 446

Qy 422 LDRLLASQSGGAPLLSTPKRERVMKMTVEEKDLEIERLTKQKELEAKQLAKAEKE 481

Db 447 LDRLLASQSGGAPLLSTPKRERVMKMTVEEKDLEIERLTKQKELEAKQLAKAEKE 506

Qy 482 NHCPTM 487

Db 507 NHCPTM 512

RESULT 4

ADQ60230

ID ADQ60230 standard; protein; 512 AA.

XX

AC ADQ60230;

XX

DT 23-SEP-2004 (first entry)

XX

DE Human microtubule motor protein #2.

XX

KW Human; microtubule motor protein; cellular proliferation disorder;

KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;

KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;

KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.

XX

OS Homo sapiens.

XX

PN US6762043-B1.

XX

PD 13-JUL-2004.

XX

PF 06-MAR-2002; 2002US-00093317.

XX

PR 20-APR-1999; 99US-00295612.

PR

PR 20-JUN-2000; 2000US-00597292.

PR

PR 28-NOV-2000; 2000US-00724224.

XX

PA (CYTO-) CYTOKINETICS INC.

XX

PI Beraud C;

XX

DR WPI; 2004-532491/51.

XX

PT New isolated microtubule motor protein, useful for screening modulators

PT for treating cellular proliferation disorders such as cancer,

PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and

PT inflammation.

XX

Fri Nov 12 12:26:27 2004

us-10-797-893-8.rag

PS	Claim 1; SEQ ID NO 4; 26pp; English.	DT	23-SEP-2004 (first entry)
XX	The invention relates to human microtubule motor proteins and the nucleic acids encoding them. The invention also relates to a method of screening for modulators of a motor protein which has microtubule stimulated ATPase activity, a method of testing for ATPase activity of microtubule motor proteins, methods to identify candidate agents that bind to a target protein or act as a modulator of the binding characteristics or biological activity of a target protein, modulators of the target protein, and methods of treating cellular proliferation disorders such as cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders and inflammation, for treating disorders associated with kinesin-like DNA binding protein (KID) and for inhibiting KID. The sequences are used for screening for modulators of motor proteins useful for treating cellular proliferation disorders such as cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders and inflammation, for treating disorders associated with KID and for inhibiting KID and for treating autoimmune diseases, arthritis, graft rejection, inflammatory bowel disease and proliferation induced after medical procedures including surgery and angioplasty. This sequence represents a human microtubule motor protein of the invention. Note: The specification states that this sequence is encoded by the nucleic acid featured as SEQ ID NO:3, but this does not appear to be the case.	XX	Homo sapiens.
XX		OS	WO2004055050-A2.
SQ	Sequence 512 AA;	XX	01-JUL-2004.
	Query Match 99.8%; Score 2467; DB 8; Length 512;	XX	10-DEC-2003; 2003WO-IB006434.
	Best Local Similarity 100.0%; Pred. No. 1.9e-223;	XX	10-DEC-2002; 2002US-0432699P.
	Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	XX	03-JUL-2003; 2003US-0485027P.
Qy	2 GRCLSKIGATRRPPARVRVAVLRPPFVDTGATAGSDPPCVRGMDSCSLIENWRNHQET 61	PT	(ENDO-) ENDOCUBE SAS.
Db	27 GRCLSKIGATRRPPARVRVAVLRPPFVDTGATAGSDPPCVRGMDSCSLIENWRNHQET 86	PA	(CNRS) CNRS CENT NAT RECH SCI.
Qy	62 LKQFDAPFYGERSTQDDIYAGSVQPIRLHLEQONASVLAYGPTGAGKTHMLGSPQPG 121	XX	Girard J, Amalric F, Roussigne M, Clouaire T;
Db	87 LKQFDAPFYGERSTQDDIYAGSVQPIRLHLEQONASVLAYGPTGAGKTHMLGSPQPG 146	XX	WPI; 2004-525034/50.
Qy	122 VIPALMDLLQLTRECAEGRPWALSVTMSYLYIYQEKVLDLDPASGDLVIREDCRGM 181	DR	N-PSDB; ADQ09241.
Db	147 VIPALMDLLQLTRECAEGRPWALSVTMSYLYIYQEKVLDLDPASGDLVIREDCRGM 206	XX	Modulating expression of a Thanatos (death)-Associated Protein (THAP) responsive gene for preventing or treating e.g. cancer or inflammation, comprises modulating the interaction of a THAP polypeptide with a nucleic acid.
Qy	182 LIPGLSQKPTSSFPADFERHFLPASRNTVGTATRLNQSRSHAVLLVKVDQRELAPFRQ 241	XX	Example 47; SEQ ID NO 425; 612pp; English.
Db	207 LIPGLSQKPTSSFPADFERHFLPASRNTVGTATRLNQSRSHAVLLVKVDQRELAPFRQ 266	CC	The present invention describes a method for modulating the expression of a thanatos (death)-associated protein (THAP) responsive gene. The method comprises modulating the interaction of a THAP-family polypeptide or its biological fragment with a nucleic acid, and so enhancing or repressing the expression of the THAP responsive gene. Also described: (1) a method of modulating the expression of a gene responsive to a THAP/chemokine complex; (2) a pharmaceutical carrier; (3) a transcription factor decoy element essentially of a THAP responsive element; (4) a cell comprising a transcription factor decoy described above; (5) methods of modulating the interaction between a nucleic acid and a THAP-family polypeptide or its biological fragment, or a nucleic acid and a THAP/chemokine complex; (6) a vector packaging cell line comprising a cell comprising a viral vector which comprises a promoter operably linked to a nucleic acid encoding a THAP-family polypeptide or its biological fragment; (7) a method of constructing a cell which expresses a recombinant THAP-family polypeptide; (8) a method of ameliorating symptoms associated with a condition mediated by a THAP/chemokine complex; (9) methods of identifying a test compound that modulates transcription at a THAP responsive element or that modulates the transport of a chemokine into the nucleus; (10) methods for reducing the symptoms associated with a condition selected from excessive or insufficient angiogenesis, inflammation, metastasis of a cancerous tissue, excessive or insufficient apoptosis, cardiovascular disease and neurodegenerative diseases; symptoms associated with a condition resulting from the activity of a chemokine or a THAP-family polypeptide in an individual; or symptoms associated with transcriptional repression or activation mediated by a THAP-family polypeptide in an individual; (11) a vector comprising a THAP responsive promoter operably linked to a nucleic acid encoding a detectable product; (12) a genetically engineered cell comprising the vector described above or that expresses a THAP-family polypeptide or its biological fragment; (13) an in vitro transcription reaction comprising a nucleic acid comprising a THAP responsive promoter, ribonucleotides and an RNA polymerase; and (14) an isolated mutant THAP-family polypeptide that does not bind to a chemokine. The pharmaceutical composition has antiangiogenic, antiinflammatory, cardiovascular,
Qy	242 REGKLYLIDLAGSDNRTGNKGLRKESGAINTSLFVLGKVVVDALNQGLPRVPYRDSKL 301	CC	
Db	267 REGKLYLIDLAGSDNRTGNKGLRKESGAINTSLFVLGKVVVDALNQGLPRVPYRDSKL 326	CC	
Qy	302 TRLLQDSLGGSAHSILIANIAPERFVLDTVSALNFAARKEVINRPTNIESLOPHALGP 361	CC	
Db	327 TRLLQDSLGGSAHSILIANIAPERFVLDTVSALNFAARKEVINRPTNIESLOPHALGP 386	CC	
Qy	362 VKLSQKELLGPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 421	CC	
Db	387 VKLSQKELLGPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 446	CC	
Qy	422 LDRLLASQGSQAPLLSTPKRERVMKMTVEKDLERLTKQKLEAKMLAKABEKE 481	CC	
Db	447 LDRLLASQGSQAPLLSTPKRERVMKMTVEKDLERLTKQKLEAKMLAKABEKE 506	CC	
Qy	482 NHCPTM 487	CC	
Db	507 NHCPTM 512	CC	
XX	RESULT 5	CC	
ADQ09240	ADQ09240 standard; protein; 665 AA.	CC	
XX	ADQ09240;	CC	
AC		CC	
XX		CC	

CC cytotatic, neuroprotective and osteopathic activities, and can be used
 CC as a THAP and THAP synthesis modulator. The composition can be used for
 CC modulating the expression of a THAP responsive gene. Modulation us useful
 CC for reducing symptoms of conditions such as excessive or insufficient
 CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
 CC or insufficient apoptosis, cardiovascular disease or neurodegenerative
 CC diseases. The present sequence is used in the exemplification of the
 CC present invention.
 XX
 XX
 XX

Sequence 665 AA;

Query Match 99.8%; Score 2467; DB 8; Length 665;
 Best Local Similarity 100.0%; Pred. No. 2.9e-223;
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVRVAVLRPFVDTAGASDPPCVRGMDSCSLEIANWRNHQET 61
 Db 26 GRCLSKIGATRRPPPARVRVAVLRPFVDTAGASDPPCVRGMDSCSLEIANWRNHQET 85
 QY 62 LKYQDFAFYGRSTQDDIYAGSVQPIRLHLEGGQNASVLAIGPTGAGKTHMLGSPQPG 121
 Db 86 LKYQDFAFYGRSTQDDIYAGSVQPIRLHLEGGQNASVLAIGPTGAGKTHMLGSPQPG 145
 QY 122 VIPRALMDLLQLTREGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 181
 Db 146 VIPRALMDLLQLTREGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 205
 QY 182 LIPGLSQKPTSSPADFERHFLPASNRNRTVGATRLNQSRSSHAVLLVKVDQERLAPFRQ 241
 Db 206 LIPGLSQKPTSSPADFERHFLPASNRNRTVGATRLNQSRSSHAVLLVKVDQERLAPFRQ 265
 QY 242 REGKLYLIDLAGSDNRRTGNKGLRLKESGAINTSFLVLGKVVDALNQGLPRVPRYRDKL 301
 Db 266 REGKLYLIDLAGSDNRRTGNKGLRLKESGAINTSFLVLGKVVDALNQGLPRVPRYRDKL 325
 QY 302 TRLLQDSLGGSAHSILANTAPERFVLDTSALNFAARKEVINRPTNESLOPHALGP 361
 Db 326 TRLLQDSLGGSAHSILANTAPERFVLDTSALNFAARKEVINRPTNESLOPHALGP 385
 QY 362 VKLSQKELLGPPPEAKRGPPEEIEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 421
 Db 386 VKLSQKELLGPPPEAKRGPPEEIEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 445
 QY 422 LDRLLASQSGQAPLLSTPKRERVMVLMKTVEEKDLEIRLTKQKELEAKMLAQKAEKE 481
 Db 446 LDRLLASQSGQAPLLSTPKRERVMVLMKTVEEKDLEIRLTKQKELEAKMLAQKAEKE 505
 QY 482 NHCEPTM 487
 Db 506 NHCEPTM 511

RESULT 6
 ADK40973
 ID ADK40973 standard; protein; 490 AA.

ADK40973;

06-MAY-2004 (first entry)

Novel human kinase protein #80.

XX cytotatic; immunomodulator; cardiant; neuroprotective; nootropic;
 KW antiparkinsonian; virucide; antibacterial; fungicide; ophthalmological;
 KW analgesic; hypotensive; immunosuppressive; kinase inhibitor; kinase;
 KW cancer; peripheral nervous system; central nervous system;
 KW Alzheimer's disease; Parkinson's disease; multiple sclerosis;
 KW amyotrophic lateral sclerosis; viral infection; prion infection;
 KW ocular disease; migraine; pain; sexual dysfunction; mood disorder;
 KW attention disorder; cognitive disorder; hypotension; hypertension;
 KW psychotic disorder; neurological disorder; dyskinesia;
 KW metabolic disorder; organ transplant rejection; enzyme.

XX

OS Homo sapiens.
 XX WO2003057841-A2.
 XX 17-JUL-2003.
 XX 31-DEC-2002; 2002WO-US041687.
 XX 31-DEC-2001; 2001US-0343169P.
 XX (GRIG/) GRIGORIEV I V.
 XX (SUDA/) SUDARSANAM S.
 XX Grigoriev IV, Sudarsanam S;
 XX WPI; 2003-587115/55.
 XX New isolated, enriched or purified nucleic acid molecule encoding a
 XX kinase polypeptide, useful for treating cancer, immune-related diseases,
 XX cardiovascular disease, brain or neuronal-associated diseases and
 XX metabolic disorders.
 XX Claim 1; SEQ ID NO 80; 491pp; English.
 XX The invention relates to novel isolated, enriched or purified nucleic acid
 XX molecules encoding a kinase polypeptide. The nucleic acid molecule
 XX comprises a sequence that: (a) encodes a kinase polypeptide; (b) is a
 XX complement of (a); (c) hybridizes under stringent conditions to (a) and
 XX encodes a naturally occurring kinase polypeptide; (d) encodes the
 XX polypeptide in (a), except that it lacks one or more, but not all, of an
 XX N-terminal domain, C-terminal catalytic domain, a catalytic domain, a C-
 XX terminal domain, a coiled-coil structure region, a spacer region and a C-
 XX terminal tail; or (e) is a complement of (d). The nucleic acid molecules,
 XX polypeptides, methods and substance are useful for treating cancers,
 XX immune-related diseases or disorders, cardiovascular disease, brain or
 XX neuronal-associated diseases, and metabolic disorders. The disorders are
 XX preferably cancers of the tissues or of hematopoietic origin, diseases of
 XX the central or peripheral nervous system, Alzheimer's disease,
 XX Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis,
 XX viral infections, infections caused by prions, infections caused by
 XX bacteria, infections caused by fungi, ocular diseases, migraines, pain,
 XX sexual dysfunction, mood disorders, attention disorders, cognition
 XX disorders, hypotension, hypertension, psychotic disorders, neurological
 XX disorders, dyskinesias, metabolic disorders and organ transplant
 XX rejection. This sequence corresponds to one of the kinase polypeptides of
 XX the invention.

Sequence 490 AA;

Query Match 93.4%; Score 2309; DB 7; Length 490;
 Best Local Similarity 100.0%; Pred. No. 1.5e-208;
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVRVAVLRPFVDTAGASDPPCVRGMDSCSLEIANWRNHQET 61
 Db 34 GRCLSKIGATRRPPPARVRVAVLRPFVDTAGASDPPCVRGMDSCSLEIANWRNHQET 93
 QY 62 LKYQDFAFYGRSTQDDIYAGSVQPIRLHLEGGQNASVLAIGPTGAGKTHMLGSPQPG 121
 Db 94 LKYQDFAFYGRSTQDDIYAGSVQPIRLHLEGGQNASVLAIGPTGAGKTHMLGSPQPG 153
 QY 122 VIPRALMDLLQLTREGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 181
 Db 154 VIPRALMDLLQLTREGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 213
 QY 182 LIPGLSQKPTSSPADFERHFLPASNRNRTVGATRLNQSRSSHAVLLVKVDQERLAPFRQ 241
 Db 214 LIPGLSQKPTSSPADFERHFLPASNRNRTVGATRLNQSRSSHAVLLVKVDQERLAPFRQ 273
 QY 242 REGKLYLIDLAGSDNRRTGNKGLRLKESGAINTSFLVLGKVVDALNQGLPRVPRYRDKL 301
 Db 274 REGKLYLIDLAGSDNRRTGNKGLRLKESGAINTSFLVLGKVVDALNQGLPRVPRYRDKL 333

Fri Nov 12 12:26:27 2004

us-10-797-893-8.rag

302	TELLQDSLGSAHSILIANIAPERFYDLTVSALNFAARKEVINRPTNESLOPHALGP	361	
334	TELLQDSLGSAHSILIANIAPERFYDLTVSALNFAARKEVINRPTNESLOPHALGP	393	
362	VKLSQKELIGPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKSSMDPAMLERLLS	421	
394	VKLSQKELIGPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKSSMDPAMLERLLS	453	
422	LDRLASQSQGAPILLSTPKRERVMVLMKTVEEKDL	456	
454	LDRLASQSQGAPILLSTPKRERVMVLMKTVEEKDL	488	
RESULT 7			
AAAB56650	ID	AAAB56650 standard; protein; 460 AA.	
AAAB56650;	AC		
13-MAR-2001 (first entry)	DT		
Human prostate cancer antigen protein sequence SEQ ID NO:1228.	DE		
Human; prostate cancer; prostate cancer antigen; detection; diagnosis; neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular; vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological; antibacterial; gene therapy; neural; immune; reproductive; renal; gastrointestinal; pulmonary; cardiovascular; proliferative disorder; wound; infectious disease.	KW		
Homo sapiens.	OS		
WO200055174-A1.	PN		
21-SEP-2000.	PD		
08-MAR-2000; 2000WO-US005988.	PF		
12-MAR-1999; 99US-0124270P.	PR		
(HUMA-) HUMAN GENOME SCI INC.	PA		
(ROSE/) ROSEN C A.	PA		
Rosen CA, Ruben SM;	PI		
WPI; 2000-587513/55.	XX		
N-PSDB; AAF15853.	DR		
Prostate cancer associated gene sequences, referred to as prostate cancer antigens, useful for treatment, prevention, and diagnosis of disorders such as prostate cancer.	PT		
Claim 11; Page 1649-1651; 2338pp; English.	PT		
AAF15566 to AAF16505 encode the human prostate cancer associated proteins, called prostate cancer antigens, given in AAB56363 to AAB57302. The prostate cancer antigens can have neuroprotective, cytostatic, cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal, nephrotropic, antiinfective, gynaecological and antibacterial activities, and can be used in gene therapy. The prostate cancer antigen polynucleotides may be used for detection of prostate cancer, chromosome identification, as chromosome markers, and for numerous other diagnostic or research purposes. The prostate cancer antigens may be used to treat disorders such as neural, immune, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, and proliferative disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to AAB57303 represent sequences used in the exemplification of the present invention	XX		
Sequence 460 AA;	SQ		
Query Match 86.6%; Score 2141; DB 3; Length 460; Best Local Similarity 95.9%; Pred. No. 9.6e-193;			
Matches 422; Conservative 1; Mismatches 17; Indels 0; Gaps 0;			
QY	2	GRCLSKI GATRRPPPARVRVAVRLPPFVDGTAGASDPFCVRGMDSCSLEIANWRNHQET	61
DB	4	GRCLSKI GATRRPPPARVRVAVRLPPFVDGTAGASDPFCVRGMDSCSLEIANWRNHQET	63
QY	62	LKYQDFAFYGERSTOODIYAGSVQPIRLHLLLEGONASVLAAYGTGAGKTHTMLGSPQPG	121
DB	64	LKYQDFAFYGERSTOODIYAGSVQPIRLHLLLEGONASVLAAYGTGAGKTHTMLGSPQPG	123
QY	122	VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDIVIREDCRNI	181
DB	124	VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDIVIREDCRNI	183
QY	182	LIPGLSQKPISSPADFERHFLPASRNRVTGATRLNQRSSSHAVLKVVDQERLAPFRQ	241
DB	184	LIPGLSQKPISSPADFERHFLPASRNRVTGATRLNQRSSSHAVLKVVDQERLAPFRQ	243
QY	242	REGKLYLIDLAGSDNRRGTGNKGLRKESGAINTSLFVLGKVVDALNQGIPRVPYRDSKL	301
DB	244	REGKLYLIDLAGSDNRRGTGNKGLRKESGAINTSLFVLGKVVDALNQGIPRVPYRDSKL	303
QY	302	TELLQDSLGSAHSILIANIAPERFYDLTVSALNFAARKEVINRPTNESLOPHALGP	361
DB	304	TELLQDSLGSAHSILIANIAPERFYDLTVSALNFAARKEVINRPTNESLOPHALGP	363
QY	362	VKLSQKELIGPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKSSMDPAMLERLLS	421
DB	364	VKLSQKELIGPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKSSMDPAMLERLLQ	423
QY	422	LDRLASQSQGAPILLSTPK 441	
DB	424	LGPSACLPGEPXGFSVEYFK 443	
RESULT 8			
ADC23342	ID	ADC23342 standard; protein; 346 AA.	
ADC23342;	AC		
18-DEC-2003 (first entry)	DT		
Human kinesin-like DNA binding protein (KID) (SeqID 6).	DE		
human; enzyme; motor domain; kinesin-like DNA binding protein; KID; cytostatic; cardiant; immunomodulator; antiinflammatory; gene therapy; cancer; hyperplasia; restenosis; cellular proliferation disorder; cardiac hypertrophy; immune disorder; inflammation.	KW		
Homo sapiens.	OS		
US6387644-B1.	PN		
14-MAY-2002.	PD		
28-NOV-2000; 2000US-00724224.	PF		
20-APR-1999; 99US-00295612.	PR		
20-JUN-2000; 2000US-00597292.	PR		
(CYTO-) CYTOKINETICS INC.	PA		
Beraud C;	PI		
WPI; 2003-706919/67.	XX		
N-PSDB; ADC23341.	DR		
Identifying a candidate agent as modulator of function of a target protein for treating cellular proliferation disorders by adding a candidate agent to a mixture of the target protein that directly/indirectly produces ADP or phosphate.	PT		
	PT		
	XX		

PS Claim 1; SEQ ID NO 6; 26pp; English.
XX This invention relates to a novel method for high throughput screening
CC systems used to identify compounds for the treatment of cellular
CC proliferation disorders. Specifically, it refers to candidate agents that
CC are capable of modulating the activity of target proteins having motor
CC domains, such that the target protein directly or indirectly produces ADP
CC or phosphate. Furthermore, this activity can be determined using
CC fluorescence or absorbance readouts. The present invention describes a
CC method that identifies modulators of the target protein, which is a
CC kinesin-like DNA binding protein (known as KID) as cytostatic, cardiant,
CC immunomodulators and antiinflammatories. Accordingly, through gene
CC therapy, they can be used for the treatment of cancer, hyperplasias,
CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This
CC polypeptide sequence is human KID protein (SeqID 6) of the invention.
XX
SQ Sequence 346 AA;

Query Match 71.5%; Score 1768; DB 7; Length 346;
Best Local Similarity 100.0%; Pred. No. 1e-157;
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGRCLSKIGATRRPPPARVAVRLRPFVDGTAGSDPPCVRGMDSCSLEIANWRNHQ 60
DB 1 MGRCLSKIGATRRPPPARVAVRLRPFVDGTAGSDPPCVRGMDSCSLEIANWRNHQ 60
QY 61 TLKQFDAPFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSP 120
DB 61 TLKQFDAPFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSP 120
QY 121 GVIPRALMDLLQLTREGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN 180
DB 121 GVIPRALMDLLQLTREGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN 180
QY 181 ILIPGLSQKPISSPADFERHFLPASRNRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR 240
DB 181 ILIPGLSQKPISSPADFERHFLPASRNRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR 240
QY 241 QREGKLYLIDLAGSDNRRRTGNKGLRKESGAINTSFLVLGKVVDALNQLGPRVPYRDSK 300
DB 241 QREGKLYLIDLAGSDNRRRTGNKGLRKESGAINTSFLVLGKVVDALNQLGPRVPYRDSK 300
QY 301 LTRLLQDSLGGSASHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
DB 301 LTRLLQDSLGGSASHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

RESULT 9
ADQ60232
ID ADQ60232 standard; protein; 346 AA.
XX
AC ADQ60232;
XX
DT 23-SEP-2004 (first entry)
XX
DE Human microtubule motor protein #3.
XX
KW Human; microtubule motor protein; cellular proliferation disorder;
KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;
KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;
KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.
XX
OS Homo sapiens.
XX
PN US6762043-B1.
XX
PD 13-JUL-2004.
XX
PF 06-MAR-2002; 2002US-00093317.
XX
PR 20-APR-1999; 99US-00295612.
PR 20-JUN-2000; 2000US-00597292.
PR 28-NOV-2000; 2000US-00724224.

XX (CYTO-) CYTOKINETICS INC.
XX Beraud C;
XX WPI; 2004-532491/51.
XX DR N-PSDB; ADQ60231.
XX
PT New isolated microtubule motor protein, useful for screening modulators
PT for treating cellular proliferation disorders such as cancer,
PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
PT inflammation.
XX
PS Claim 1; SEQ ID NO 6; 26pp; English.
XX
CC The invention relates to human microtubule motor proteins and the nucleic
CC acids encoding them. The invention also relates to a method of screening
CC for modulators of a motor protein which has microtubule stimulated ATPase
CC activity, a method of testing for ATPase activity of microtubule motor
CC proteins, methods to identify candidate agents that bind to a target
CC protein or act as a modulator of the binding characteristics or
CC biological activity of a target protein, modulators of the target
CC protein, and methods of treating cellular proliferation disorders such as
CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
CC and inflammation, for treating disorders associated with kinesin-like DNA
CC binding protein (KID) and for inhibiting KID. The sequences are used for
CC screening for modulators of motor proteins useful for treating cellular
CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
CC hypertrophy, immune disorders and inflammation, for treating autoimmune
CC diseases, arthritis, graft rejection, inflammatory bowel disease and
CC proliferation induced after medical procedures including surgery and
CC angioplasty. This sequence represents a human microtubule motor protein
CC of the invention.
XX
SQ Sequence 346 AA;

Query Match 71.5%; Score 1768; DB 8; Length 346;
Best Local Similarity 100.0%; Pred. No. 1e-157;
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGRCLSKIGATRRPPPARVAVRLRPFVDGTAGSDPPCVRGMDSCSLEIANWRNHQ 60
DB 1 MGRCLSKIGATRRPPPARVAVRLRPFVDGTAGSDPPCVRGMDSCSLEIANWRNHQ 60
QY 61 TLKQFDAPFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSP 120
DB 61 TLKQFDAPFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSP 120
QY 121 GVIPRALMDLLQLTREGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN 180
DB 121 GVIPRALMDLLQLTREGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN 180
QY 181 ILIPGLSQKPISSPADFERHFLPASRNRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR 240
DB 181 ILIPGLSQKPISSPADFERHFLPASRNRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR 240
QY 241 QREGKLYLIDLAGSDNRRRTGNKGLRKESGAINTSFLVLGKVVDALNQLGPRVPYRDSK 300
DB 241 QREGKLYLIDLAGSDNRRRTGNKGLRKESGAINTSFLVLGKVVDALNQLGPRVPYRDSK 300
QY 301 LTRLLQDSLGGSASHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
DB 301 LTRLLQDSLGGSASHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

RESULT 10
ADQ23338
ID ADQ23338 standard; protein; 370 AA.
XX
AC ADQ23338;
XX
DT 18-DEC-2003 (first entry)

XX DE Human kinesin-like DNA binding protein (KID) (SeqID 2).

XX KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID;

XX KW cytoskeletal; cardiant; immunomodulator; antiinflammatory; gene therapy;

XX KW cancer; hyperplasia; restenosis; cellular proliferation disorder;

XX KW cardiac hypertrophy; immune disorder; inflammation.

XX OS Homo sapiens.

XX XX

XX FT Key .Location/Qualifiers

XX FT Misc-difference 1

XX FT /note= "Encoded by ATGCA"

XX XX

XX PN US6387644-B1.

XX XX

XX PD 14-MAY-2002.

XX XX

XX PF 28-NOV-2000; 2000US-00724224.

XX XX

XX PR 20-APR-1999; 99US-00295612.

XX PR 20-JUN-2000; 2000US-00597292.

XX XX

XX PA (CYTO-) CYTOKINETICS INC.

XX XX

XX PI Beraud C;

XX XX

XX DR WPI; 2003-706919/67.

XX DR N-PSDB; ADC23337.

XX XX

XX PT Identifying a candidate agent as modulator of function of a target

XX PT protein for treating cellular proliferation disorders by adding a

XX PT candidate agent to a mixture of the target protein that

XX PT directly/indirectly produces ADP or phosphate.

XX XX

XX PS Claim 1; SEQ ID NO 2; 26pp; English.

XX XX

XX CC This invention relates to a novel method for high throughput screening

XX CC systems used to identify compounds for the treatment of cellular

XX CC proliferation disorders. Specifically, it refers to candidate agents that

XX CC are capable of modulating the activity of target proteins having motor

XX CC domains, such that the target protein directly or indirectly produces ADP

XX CC or phosphate. Furthermore, this activity can be determined using

XX CC fluorescence or absorbance readouts. The present invention describes a

XX CC method that identifies modulators of the target protein, which is a

XX CC kinesin-like DNA binding protein (known as KID) as cytoskeletal, cardiant,

XX CC immunomodulators and antiinflammatory. Accordingly, through gene

XX CC therapy, they can be used for the treatment of cancer, hyperplasias,

XX CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This

XX CC polypeptide sequence is human KID protein (SeqID 2) of the invention.

XX XX

XX SQ Sequence 370 AA;

Query Match 71.3%; Score 1763; DB 7; Length 370;

Best Local Similarity 100.0%; Pred. No. 3.3e-157;

Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GCRLSKIGATRRPPPPARVAVRLRPVDTAGASDPPCVRGWDCSLEIANRNHOET 61

DB 26 GCRLSKIGATRRPPPPARVAVRLRPVDTAGASDPPCVRGWDCSLEIANRNHOET 85

QY 62 LKYQDFAYGERSTQDDIYAGSVQPIRLHLEGGQNASVLAYGPTGAGKTHMLGSPQPG 121

DB 86 LKYQDFAYGERSTQDDIYAGSVQPIRLHLEGGQNASVLAYGPTGAGKTHMLGSPQPG 145

QY 122 VIPRALMDLLQUTREEGAGRPWALSVTMSYLEIQEKVLDLLDPASGDLVIREDCRGN 181

DB 146 VIPRALMDLLQUTREEGAGRPWALSVTMSYLEIQEKVLDLLDPASGDLVIREDCRGN 205

QY 182 LIPGLSQPISSFADFERHFLPASRNRTVGATRLNQRSSRHVALLVKVDQREIAPFRQ 241

DB 206 LIPGLSQPISSFADFERHFLPASRNRTVGATRLNQRSSRHVALLVKVDQREIAPFRQ 265

QY 242 REKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVKGKVVDAALNQGILPRVYRDSKL 301

DB 266 REKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVKGKVVDAALNQGILPRVYRDSKL 325

QY 302 TRLLQSLGGSASISILIANIAPERRFYLDVTVSALNFAARSKEVIN 346

DB 326 TRLLQSLGGSASISILIANIAPERRFYLDVTVSALNFAARSKEVIN 370

RESULT 11

ADO60228

ID ADO60228 standard; protein; 370 AA.

XX AC ADO60228;

XX DT 23-SEP-2004 (first entry)

XX XX Human microtubule motor protein #1.

XX KW Human; microtubule motor protein; cellular proliferation disorder;

XX KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;

XX KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;

XX KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.

XX OS Homo sapiens.

XX XX

XX PN US6762043-B1.

XX PD 13-JUL-2004.

XX XX

XX PF 06-MAR-2002; 2002US-00093317.

XX PR 20-APR-1999; 99US-00295612.

XX PR 20-JUN-2000; 2000US-00597292.

XX PR 28-NOV-2000; 2000US-00724224.

XX XX

XX PA (CYTO-) CYTOKINETICS INC.

XX XX

XX PI Beraud C;

XX XX

XX DR WPI; 2004-532491/51.

XX XX

XX PT New isolated microtubule motor protein, useful for screening modulators

XX PT for treating cellular proliferation disorders such as cancer,

XX PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and

XX PT inflammation..

XX XX

XX PS Claim 1; SEQ ID NO 2; 26pp; English.

XX XX

XX CC The invention relates to human microtubule motor proteins and the nucleic

XX CC acids encoding them. The invention also relates to a method of screening

XX CC for modulators of a motor protein which has microtubule stimulated ATPase

XX CC activity, a method of testing for ATPase activity of microtubule motor

XX CC proteins, methods to identify candidate agents that bind to a target

XX CC protein or act as a modulator of the binding characteristics or

XX CC biological activity of a target protein, modulators of the target

XX CC protein, and methods of treating cellular proliferation disorders such as

XX CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders

XX CC and inflammation, for treating disorders associated with kinesin-like DNA

XX CC binding protein (KID) and for inhibiting KID. The sequences are used for

XX CC screening for modulators of motor proteins useful for treating cellular

XX CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac

XX CC hypertrophy, immune disorders and inflammation, for treating disorders

XX CC associated with KID and for inhibiting KID and for treating autoimmune

XX CC diseases, arthritis, graft rejection, inflammatory bowel disease and

XX CC proliferation induced after medical procedures including surgery and

XX CC angioplasty. This sequence represents a human microtubule motor protein

XX CC of the invention. Note: The specification states that this sequence is

XX CC encoded by the nucleic acid featured as SEQ ID NO:1, but this does not

XX CC appear to be the case.

XX XX

XX SQ Sequence 370 AA;

Query Match 71.3%; Score 1763; DB 8; Length 370;
Best Local Similarity 100.0%; Pred. No. 3.3e-157;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPARVAVRLRPVDTAGASDPPCVRGNDSCSLEIANWRNHOET 61
DB 26 GRCLSKIGATRRPPARVAVRLRPVDTAGASDPPCVRGNDSCSLEIANWRNHOET 85
QY 62 LKYQDFAFYGRSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 121
DB 86 LKYQDFAFYGRSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 145
QY 122 VIPALMDLLOLTREBAGRPWALSVMTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 181
DB 146 VIPALMDLLOLTREBAGRPWALSVMTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 205
QY 182 LIPGLSKPISFPADFERHFLPASNETVGATRLNQRSSRSRSHAVLLVKVDORERLAPFRQ 241
DB 206 LIPGLSKPISFPADFERHFLPASNETVGATRLNQRSSRSRSHAVLLVKVDORERLAPFRQ 265
QY 242 REGKLYLIDLAGSDNRRTGNKGLRLKESGAINTSFLVLGKVDALNQGRLPRVPYRDSKL 301
DB 266 REGKLYLIDLAGSDNRRTGNKGLRLKESGAINTSFLVLGKVDALNQGRLPRVPYRDSKL 325
QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
DB 326 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 370

RESULT 12

ABB71112
ID ABB71112 standard; protein; 784 AA.
AC ABB71112;
XX
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 40128.
DE Drosophila melanogaster polypeptide SEQ ID NO 40128.
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
OS Drosophila melanogaster.
XX
XX W0200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US0009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
DR N-PSDB; ABL15215.
XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 40128; 21pp + Sequence Listing; English.
XX

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC

CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

QY Sequence 784 AA;

Query Match 24.9%; Score 616; DB 4; Length 784;
Best Local Similarity 33.1%; Pred. No. 2.1e-48;
Matches 176; Conservative 84; Mismatches 183; Indels 88; Gaps 18;

QY 12 TRRP-----PPARVAVRLRPVDTAGASDPPCV-----RGMDSCSLEIANW--RN 57
DB 5 SRPFGTSSQTPNECVVCRPMNSRERSPEVNVYVNRGV----VELQVVDGN 60
QY 58 HOETLYQDFAFYGRSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHM---L 114
DB 61 KEQKVFYTDAAVDASATQTLTLYHEVFPVPLVSSVLEGFNGCIFYGOTGTGKTFTMGVR 120
QY 115 GSPEQPGVIPRALMDL-LQLTREBAGRPWALSVMTMSYLEIYQEKVLDLDPASGDLVI 173
DB 121 GNDELMLGIIIPRTFEQIWLHINRTEN-----FQFLVDVSYLEIYMEELRDLKPNKSHLEV 175
QY 174 REDCRGN-ILIPGLSKPISFPADFERHFLPASNETVGATRLNQRSSRSRSHAVLLVKVDQ 232
DB 176 RE--RGSVYVPLNLHAINCKSVEDMIKVMQGNKRTVGFTNMNHSRSHAFMIKIE 233
QY 233 RERLAPPRQREGKLYLIDLAGSDNRRTGNKGLRLKESGAINTSFLVLGKVDALNQGRL 292
DB 234 CD-TEYNTIKVGNLIDLAGSROSKTGASERLKEASKINLALSLSGNVISALAESSP 292
QY 293 RVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPTNE 352
DB 293 HVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPTNE 352
QY 353 SLQPHALGPVKLSQ-----KELGPPPEAKRA-----ROPEEEI----- 386
DB 353 DPQ-----DAKLEYQBEIERLRLKLGPPQQORSEKQVAKQKRVKPKETVTTKEMSDS 407
QY 387 -----GSPEPMAAPASASQKLSPLQKLSMDPAMLERLLSLDRLLASQSSQ 433
DB 408 LQVSTIEQPVEDSDPE--GAESSEDKENAEVAKSNEELEERVEN-SKLAALKLALE 463
QY 434 APILLSTPKRERMVLMKTVBEKLEIER-----LTKKQKLEAKMLAQKASE 479
DB 464 GLVRRGKN---LLDTYSERQIELEKLVIEAKRKKREIEIQOQLQEE 510

RESULT 13

ABB65183
ID ABB65183 standard; protein; 677 AA.
XX
XX ABB65183;
AC ABB65183;
XX
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 22341.
DE Drosophila melanogaster polypeptide SEQ ID NO 22341.
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
OS Drosophila melanogaster.
XX W0200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US0009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
XX

PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
DR N-PSDB; ABL09286.
XX
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 22341; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 677 AA;

Query Match 24.0%; Score 593; DB 4; Length 677;
Best Local Similarity 33.9%; Pred. No. 2.5e-46;
Matches 167; Conservative 78; Mismatches 190; Indels 58; Gaps 12;

QY 20 VRVAVLRPFVDG--TAGADPPCVRGMD--SCSLEITANRNHQTLYQDFAFYGRSTQ 76
DB 21 VRVVTRPMDKNLSAGALSISVDKINRAITVMPNATANEPPKTYTFDVFQGSNQ 80

QY 77 QDIYAGSVQPIRLHLEGQNASLAYGPTGAGKTHMLGSPEQP---GVIPRALMDL--- 130
DB 81 MDLYVDTARPIVQKVLGYNGTILAYGQTGTGKTYTMSGNPDPSPQTKGIIPNFAHIFGH 140

QY 131 LQLTREBAGRPWALSVTWYSLEYIQEVKVLDDLPASG-DLVIREDCRGNLILPGLSQ 189
DB 141 IAKAKEN-----QKFLRVSYMEIYNEEVRDLLGKDVGKSLVKRDPDVGIVFKDLSGY 194

QY 190 PISSEADFERHFLPASNRRTVGTALNRSSRHAVLLVKVDQERLAPFRQ--REGKLY 247
DB 195 VHNWADLENIMRLGNKRAVGTAKNQESSRSHAIFSIIVERSELGEGDVQHVVRMGKLO 254

QY 248 LIDLAGSEDRRTNGKLRKESGAINISLFLVGLKVVDALNOG-LPRVPRDSKLTLLQ 306
DB 255 LVDLAGSERQSKTQASGQRLKEATKINLSLVGLNVISALVDGKSTHPIYRNSKLTLLQ 314

QY 307 DSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPTNE----SLQPHALGPV 362
DB 315 DSLGNSKTVNCATISPADSNYMETITSLRYASRAKNIQRMINEEPKDALLRHQEEI 374

QY 363 KLSQKEL-LGPPEAKRARGPEEIEIGSPPEMAFPASASQKSLPQLKSSMDPAMLERLLS 421
DB 375 ARLRKQLEEGDSLEEPSSEEDTADDELEAPLEIELESSTTQAVEKPKKREK--- 431

QY 422 LDRLLASQGSOGAPLLSTPKRERVMKTVVEKDLERLKTQKELEAKM----- 472
DB 432 -----TDAEKEELAKRNEHOKETIHAKEQETLRNKLIVSLGKILV 473

QY 473 ----LAOKAEKE 481
DB 474 GGNLLEKAQIQE 486

RESULT 14
ADK40926
ID ADK40926 standard; protein; 1237 AA.
XX
AC ADK40926;
XX
DT 06-MAY-2004 (first entry)
XX

DE Novel human kinase protein #33.
XX
XX cytostatic; immunomodulator; cardiant; neuroprotective; nootropic;
KW antiparkinsonian; virucide; antibacterial; fungicide; ophthalmological;
KW analgesic; hypotensive; immunosuppressive; kinase inhibitor; kinase;
KW cancer; peripheral nervous system; disease; multiple sclerosis;
KW Alzheimer's disease; Parkinson's disease; viral infection; prion infection;
KW amyotrophic lateral sclerosis; sexual dysfunction; mood disorder;
KW ocular disease; migraine; pain; sexual dysfunction; hypertension;
KW attention disorder; cognition disorder; hypotension; hypertension;
KW psychotic disorder; neurological disorder; dyskinesia;
KW metabolic disorder; organ transplant rejection; enzyme.
XX
XX Homo sapiens.
XX WO2003057841-A2.
XX 17-JUL-2003.
XX
XX 31-DEC-2002; 2002WO-US041687.
XX 31-DEC-2001; 2001US-0343169P.
XX (GRIG/) GRIGORIEV I V.
XX (SUDA/) SUDARSANAM S.
XX Grigoriev IV, Sudarsanam S;
XX WPI; 2003-587115/55.
XX
XX New isolated, enriched or purified nucleic acid molecule encoding a
PT kinase polypeptide, useful for treating cancer, immune-related diseases,
PT cardiovascular disease, brain or neuronal-associated diseases and
PT metabolic disorders.
XX
XX Claim 1; SEQ ID NO 33; 491pp; English.
XX
XX The invention relates to novel isolated, enriched or purified nucleic acid
CC molecules encoding a kinase polypeptide. The nucleic acid molecule
CC comprises a sequence that: (a) encodes a kinase polypeptide; (b) is a
CC complement of (a); (c) hybridizes under stringent conditions to (a) and
CC encodes a naturally occurring kinase polypeptide; (d) encodes the
CC polypeptide in (a), except that it lacks one or more, but not all, of an
CC N-terminal domain, C-terminal catalytic domain, a spacer region and a C-
CC terminal tail; or (e) is a complement of (d). The nucleic acid molecules,
CC polypeptides, methods and substance are useful for treating cancer,
CC immune-related diseases or disorders, cardiovascular disease, brain or
CC neuronal-associated diseases, and metabolic disorders. The disorders are
CC preferably cancers of the tissues or of hematopoietic origin, diseases of
CC the central or peripheral nervous system, Alzheimer's disease,
CC Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis,
CC viral infections, infections caused by prions, infections caused by
CC bacteria, infections caused by fungi, ocular diseases, migraines, pain,
CC sexual dysfunction, mood disorders, attention disorders, cognition
CC disorders, hypotension, hypertension, psychotic disorders, neurological
CC disorders, dyskinesias, metabolic disorders and organ transplant
CC rejection. This sequence corresponds to one of the kinase polypeptides of
CC the invention.
XX
XX SQ Sequence 1237 AA;

Query Match 23.4%; Score 578.5; DB 7; Length 1237;
Best Local Similarity 32.4%; Pred. No. 1.5e-44;
Matches 157; Conservative 97; Mismatches 172; Indels 59; Gaps 13;

QY 20 VRVAVLRPFVDGTAGADPPC-----VRGWDSCSLEITANRNHQTLYQDFAF 69
DB 13 VRVALRCRLVPKEISEGQCMCLSFVPGTQVVVGIDK-----SFYDFV 57

QY 70 YGERSTQODIYAGSVQPIRLHLEGQNASLAYGPTGAGKTHMLGS-----PEQP--GV 122
DB 58 FDPCTQEVEVFNKAVAPLIKGFKNATVLAQYQTSKTKYSGMGAYTAQEENPTVGI 117

Db 391 KNOQLVEENEKLSRGLSEAGQTAQML-----ERILLTEQANEKNKNAKLEELRQHACKL 445

Qy 469 EAKMLAQKAEKE 481

Db : : : : :
446 DLQKLVTLEDOE 458

Search completed: November 5, 2004, 18:44:42
Job time : 101.956 secs

GenCore version 5.1.6.
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 04:25:02 ; Search time 3876.13 Seconds
(without alignments)
10482.189 Million cell updates/sec

Title: US-10-797-893-1
Perfect score: 1115
Sequence: 1 atgcagccgaggcggtctg.....ccaaggagtgatcaattga 1115

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gse1:*

9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1107.4	99.3	2057	3	CR615781 full-length
2	1093.4	98.1	2025	3	CR590974 full-length
3	1045.4	93.8	1124	1	AL545712 full-length
4	1043.4	93.6	2016	3	CR590954 full-length
5	1002.8	89.9	1027	5	CR337406 full-length
6	980.8	88.0	1083	1	AL559772 full-length
7	978.6	87.8	1032	5	CR337353 full-length
8	942	84.5	959	5	CR337358 full-length
9	934	83.8	958	5	CR354166 full-length
10	921.4	82.6	999	1	AL559769 full-length
11	921.4	82.6	1094	5	CR376199 full-length
12	917	82.2	1031	4	BM910587 full-length
13	909	81.5	1019	4	BM546877 full-length
14	907.8	81.4	978	5	CR400806 full-length
15	905	81.2	1016	1	AL518361 full-length
16	889.8	79.8	1051	4	BM806996 full-length
17	889.6	79.8	980	5	CR404364 full-length
18	883.2	79.7	1041	4	BM478215 full-length
19	887.4	79.6	915	5	CR383181 full-length
20	862.4	77.3	926	1	AL578338 full-length
21	861.8	77.3	1040	4	BM546897 full-length
22	861.6	77.3	991	5	CR333336 full-length
23	851.8	76.4	865	1	AL555550 full-length
24	848.8	76.1	865	6	CD109092 full-length

25	844	75.7	990	5	BX399192	BX399192
26	836.4	75.0	1139	4	BM553107	BM553107
27	833.6	74.8	971	5	BQ53671	BQ53671
28	833	74.7	838	5	BX341048	BX341048
29	831.8	74.6	851	5	BUI88590	BUI88590
30	831.6	74.6	942	5	BUS27647	BUS27647
31	823	73.8	995	1	AL559752	AL559752
32	821.8	73.7	846	5	BX404325	BX404325
33	819	73.5	921	5	BUS60011	BUS60011
34	816.4	73.2	878	1	AL576578	AL576578
35	815.2	73.1	872	5	BUS21359	BUS21359
36	808.6	72.5	873	5	BQ955064	BQ955064
37	808.6	72.5	971	5	CR397059	CR397059
38	801.2	71.9	880	4	BG683241	BG683241
39	801	71.8	849	5	BX343522	BX343522
40	789	70.8	961	5	BUI66881	BUI66881
41	787	70.6	813	4	BI835505	BI835505
42	786.6	70.5	911	5	BUS31772	BUS31772
43	786.4	70.5	956	1	AL558200	AL558200
44	781.4	70.1	806	4	BI915916	BI915916
45	781.4	70.1	913	5	BX445789	BX445789

ALIGNMENTS

RESULT 1
LOCUS CR615781 2057 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CSODK012YE23 of HeLa cells Cot 25-normalized of Homo sapiens (human).
ACCESSION CR615781
VERSION CR615781.1 GI:50496588
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2057)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InvitroGen Corporation 1600 Faraday Avenue
REFERENCE 2 (bases 1 to 2057)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES
source location/Qualifiers
1..2057
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODK012YE23"
/tissue_type="HeLa cells Cot 25-normalized"
/plasmid="pCMVSPORT_6"

Query Match 99.3%; Score 1107.4; DB 3; Length 2057;
Best Local Similarity 99.9%; Pred. No. 3.3e-294;
Matches 1108; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 GCCCGGGCGCTCGACGACGACGACGACGACGATGGCGGAGCTTCAGCGCGCGG 65

Db 7 GCCCGGGCGCTCGACGACGACGACGACGACGATGGCGGAGCTTCAGCGCGCGG 66

440 GATCCCGGGCTCTCATGACCTCTCTGAGCTCAACAGGAGGAGGGTCCGAGGGCG 499
 421 GATCCCGGGCTCTCATGACCTCTCTGAGCTCAACAGGAGGAGGGTCCGAGGGCG 480
 500 GCATGGGGCCCTTCTGTCACCATCTTACCTAGAGATCTACAGGAGAGGATATTAGA 559
 481 GCATGGGGCCCTTCTGTCACCATCTTACCTAGAGATCTACAGGAGAGGATATTAGA 540
 560 CTTCTCTGAGCCCTCTCTGAGAGACCTGCTTAATCCGAGAGACTGCGGGGGGAATATCT 619
 541 CTTCTCTGAGCCCTCTCTGAGAGACCTGCTTAATCCGAGAGACTGCGGGGGGAATATCT 600
 620 GATTCGGGGTCTCTCCAGAGCCCATAGTACTGTTGCTGATTTTGGAGCGGACTTCT 679
 601 GATTCGGGGTCTCTCCAGAGCCCATAGTACTGTTGCTGATTTTGGAGCGGACTTCT 660
 680 GCAGGCCAGTCAAAATCGAGCTGTAGGAGCCACCGGCTCAACAGGCGTCTCTCCGGAG 739
 661 GCAGGCCAGTCAAAATCGAGCTGTAGGAGCCACCGGCTCAACAGGCGTCTCTCCGGAG 720
 740 TCATGCTGTGCTCTCTGCTCAAGTGGAGCGGGAACTTTGGCCCAATTTTGGCCAGCG 799
 721 TCATGCTGTGCTCTCTGCTCAAGTGGAGCGGGAACTTTGGCCCAATTTTGGCCAGCG 780
 800 AGAGGGAATCTTACCTGATTTGACTTGGCTGGGTGAGGCAACACCGGCGCACAGGCAA 859
 781 AGAGGGAATCTTACCTGATTTGACTTGGCTGGGTGAGGCAACACCGGCGCACAGGCAA 840
 860 CAAGGGCTTCCGCTAAAGAGAGTGGAGGCAATCAACCTCTGTTTGTCTGGGCAA 919
 841 CAAGGGCTTCCGCTAAAGAGAGTGGAGGCAATCAACCTCTGTTTGTCTGGGCAA 900
 920 AGTGTGTAGTGGCTGAATCAGGGCTCCCTGCTGTACCTTATCGGACAGCAAGCTCAC 979
 901 AGTGTGTAGTGGCTGAATCAGGGCTCCCTGCTGTACCTTATCGGACAGCAAGCTCAC 960
 980 TCGCTATTGAGGAGCTCTCTGGGTGGCTCAGCCACAGTATCTTATGCAACATTGC 1039
 961 TCGCTATTGAGGAGCTCTCTGGGTGGCTCAGCCACAGTATCTTATGCAACATTGC 1020
 1040 CCTCAGAGAGCCTTCTACCTAGACACAGTCTCCGACATCAATCTTGTCCAGGTCCAA 1099
 1021 CCTCAGAGAGCCTTCTACCTAGACACAGTCTCCGACATCAATCTTGTCCAGGTCCAA 1080
 1100 GGAGGTGATCAATTG 1114
 1081 GGAGGTGATCAATCG 1095

RESULT 3
 LOCUS AL545712 1124 bp mRNA linear EST 25-MAR-2004
 DEFINITION AL545712 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DI009YJ01 5-PRIME, mRNA sequence.
 ACCESSION AL545712
 VERSION AL545712.3 GI:45746191
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1124)
 Li W.B., Gruber C., Jessee J. and Polayes D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:31267547.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7766.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?e=CS0DI009CE01QP1&c=7766.r.
 Location/Qualifiers
 1..1124
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DI009YJ01"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
 source

ORIGIN

Query Match 93.8%; Score 1045.4; DB 1; Length 1124;
 Best Local Similarity 96.6%; Pred No. 3.8e-277;
 Matches 1059; Conservative 24; Mismatches 11; Indels 2; Gaps 2;
 QY 20 GACGACGAGGCGGACGAGATGGCGGAGCTTTCAGCGGCGGCGATCTCAGGAGCTCG 79
 Db 1 GACGACGAGGCGGACGAGATGGCGGAGCTTTCAGCGGCGGCGATCTCAGGAGCTCG 60
 QY 80 TCGTGTGGCTTAAGCAAGATTTGGAGTACTCGTGTCTCAGCTCGGCTGTAAGGCT 139
 Db 61 TCGTGTGGCTTAAGCAAGATTTGGAGTACTCGTGTCTCAGCTCGGCTGTAAGGCT 120
 QY 140 GCGTGTGGCTTAAGCAAGATTTGGAGTACTCGTGTCTCAGCTCGGCTGTAAGGCT 199
 Db 121 GCGTGTGGCTTAAGCAAGATTTGGAGTACTCGTGTCTCAGCTCGGCTGTAAGGCT 180
 QY 200 GCGGCGCATGACAGCTGCTCTAGAGATTTGCTAACTGGAGGAACACACAGGAGACTCT 259
 Db 181 GCGGCGCATGACAGCTGCTCTAGAGATTTGCTAACTGGAGGAACACACAGGAGACTCT 240
 QY 260 CAATACAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGAGGAGACATCTATGACGG 319
 Db 241 CAATACAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGAGGAGACATCTATGACGG 300
 QY 320 TTCACTGAGCCCATCTCAAGGAGTCTGCTGAAGGAGGAGATGCTGCTGCTGCTGCTGCT 379
 Db 301 TTCACTGAGCCCATCTCAAGGAGTCTGCTGAAGGAGGAGATGCTGCTGCTGCTGCTGCT 360
 QY 380 TGGACCCACAGGAGCTGGGAGAGCGACACAACTGCTGGGAGGAGGAGGAGGAGGAGG 439
 Db 361 TGGACCCACAGGAGCTGGGAGAGCGACACAACTGCTGGGAGGAGGAGGAGGAGGAGG 420
 QY 440 GATCCCGGGGCTCTCATGAGCTCTCTGAGCTCAACAGGAGGAGGAGGAGGAGGAGG 499
 Db 421 GATCCCGGGGCTCTCATGAGCTCTCTGAGCTCAACAGGAGGAGGAGGAGGAGGAGG 480
 QY 500 GCCATGGGCGCTTCTGCTCAGCATGCTTACTAGAGATCTACAGGAGAGGAGGATATTAGA 559
 Db 481 GCCATGGGCGCTTCTGCTCAGCATGCTTACTAGAGATCTACAGGAGAGGAGGATATTAGA 540
 QY 560 CTTCTCTGAGCCCTCTCTGAGAGACCTGCTTAATCCGAGAGACTGCGGGGGGAATATCT 619
 Db 541 CTTCTCTGAGCCCTCTCTGAGAGACCTGCTTAATCCGAGAGACTGCGGGGGGAATATCT 600
 QY 620 GATTCGGGGTCTCTCCAGAGCCCATAGTACTGTTGCTGATTTTGGAGCGGACTTCT 679
 Db 601 GATTCGGGGTCTCTCCAGAGCCCATAGTACTGTTGCTGATTTTGGAGCGGACTTCT 660
 QY 680 GCAGGCCAGTCAAAATCGAGCTGTAGGAGCCACCGGCTCAACAGGCGTCTCTCCGGAG 739
 Db 661 GCAGGCCAGTCAAAATCGAGCTGTAGGAGCCACCGGCTCAACAGGCGTCTCTCCGGAG 720
 QY 740 TCATGCTGTGCTCTCTGCTCAAGTGGAGCGGGAACTTTGGCCCAATTTTGGCCAGCG 799

Db 721 TCATGCTGTGCTCCTGGTCAAGGKGGACCGGGGACAGTTTGGCCCAATTTGCCAGCG 780

Qy 800 AGAGGGAAAACCTTACCTGATTGACTTGGCTGGGTGAGAGCAACCGGCGCA CAGGCAA 859

Db 781 AGAGGGAAAACCTTACCTGATTGACTTGGCTGGGTGAGAGCAACCGGCGCA CAGGCAA 840

Qy 860 CAAGGGCTTCGGCT-AAAAGAGAGTGAGGCAATCAACACCTCCCTGTTTGTCTGGGCA 918

Db 841 CAAGGGCTTCGGSTAAAAGAGAGTGAGGCAATCAACACCTCCCTGTTTGTCTGGGCA 900

Qy 919 AAGTGGTAGATGCGCTGAATCAGGCGCTCCCTCGTGTACCTTATCGGAGCAGCAAGCTCA 978

Db 901 AAGTGGTAGATGCGCTGAATCAGGCGCTCCCTCGTGTACCTTATCGGRCAGCAAGCTCA 960

Qy 979 CTCGCCATTGACAGACTCTCGGCTGGGTGAGGCGGAGAGTATCTTATTGCGCAATTG 1038

Db 961 CTCGCCATTGACAGGCTCTCTGGGKGSTCAG-CCACAGTATCTTATTGCCAACATTG 1019

Qy 1039 CCCCTGAGAGCGCTTCTTACCTAGACACAGTCTCCGCACTCAACTTTTGTGCGCAGGTCCA 1098

Db 1020 CCCCTGAGAGCGCTTCTTACCTAGACACAGTCTCCGCACTCAACTTTTGTGCGAGGCCA 1079

Qy 1099 AGGAGGTGATCAATTG 1114

Db 1080 AGGGGKGRCTCAATCG 1095

RESULT 4

CR590954

LOCUS

DEFINITION

2016 bp mRNA linear HTC 21-JUL-2004

full-length cDNA clone CS0D1040Y106 of Placenta Cot 25-normalized

of Homo sapiens (human).

ACCESSION

CR590954

VERSION

CR590954.1 Gi:50471761

KEYWORDS

HTC; CNSLT cdNA.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 2016)

Li, W.B., Gruber, C., Jesses, J. and Polayes, D.

Full-length cDNA libraries and normalization.

Unpublished

Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue

2 (bases 1 to 2016)

Genoscope.

Direct Submission

Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

1st enriched cDNA was primed with a NotI-oligo(dT) primer. Five prime end strand, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES

source

1. 2016

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0D1040Y106"

/tissue_type="Placenta Cot 25-normalized"

/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 93.6%; Score 1043.4; DB 3; Length 2016;

Best Local Similarity 99.9%; Pred. No. 1.5e-276;

Matches 1044; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 70 CAGGAGCTGGTGGCTGAAGCAAGATTGGAGCTACTGCTGCTCCACCTCCAGCTC 129

ORIGIN	Query Match	88.0%;	Score 980.8;	DB 1;	Length 1083;
	Best Local Similarity	96.4%;	Pred. No. 2.5e-259;	Indels	Gaps
	Matches 1010;	Conservative	13;	Mismatches 21;	
Qy	10	CGGCGGGCTCGACGCGAGGCGGACGCGAGATGGCGGACGTTCAAGCGCGCGGATCT	69		
Db	1	CGGCGGGCTCGACGCGAGGCGGACGCGAGATGGCGGACGTTCAAGCGCGCGGATCT	60		
Qy	70	CAGAGCTGGTGGCTGGCTGGCTAGCAGATTTGGAGCTACTCGTCCTCACTCCAGCTC	129		
Db	61	CAGAGCTGGTGGCTGGCTGGCTAGCAGATTTGGAGCTACTCGTCCTCACTCCAGCTC	120		
Qy	130	CGCTAAGGGTGGCTGGCGACTGGCGCCATTGTGGATGGGAACAGCGGGAGCAAGTGATC	189		
Db	121	CGCTAAGGGTGGCTGGCGACTGGCGCCATTGTGGATGGGAACAGCGGGAGCAAGTGATC	180		
Qy	190	CCCCCTGTGGGGGATGCAAGCTGCTCTTAGAGATTGCTAACTGGAGGAACCAACC	249		
Db	181	CCCCCTGTGGGGGATGCAAGCTGCTCTTAGAGATTGCTAACTGGAGGAACCAACC	240		
Qy	250	AGGAGACTCTCAAAATACCAAGTTTGTATGGCTTCTATGGGGAGAGGATGACTCAGCAGGACA	309		
Db	241	AGGAGACTCTCAAAATACCAAGTTTGTATGGCTTCTATGGGGAGAGGATGACTCAGCAGGACA	300		
Qy	310	TCTATGCAGGTTCACTGAGCGCCATCTTAAGGCATCTGTGGAAGCGGAGAGTAATGCCAGTG	369		
Db	301	TCTATGCAGGTTCACTGAGCGCCATCTTAAGGCATCTGTGGAAGCGGAGAGTAATGCCAGTG	360		
Qy	370	TGCTTGGCTATGGACCCACAGGAGCTGGGAAGACGCAACAATGTGGGCGACCCAGAGC	429		
Db	361	TGCTTGGCTATGGACCCACAGGAGCTGGGAAGACGCAACAATGTGGGCGACCCAGAGC	420		
Qy	430	AACCTGGGGTGAATCCGGGGCTCTCATGGACCTCTGCAGCTCAACAGGAGGAGGGTG	489		
Db	421	AACCTGGGGTGAATCCGGGGCTCTCATGGACCTCTGCAGCTCAACAGGAGGAGGGTG	480		
Qy	490	CCGAGGGCGGCGATCGGGCCCTTCTGTCCACATGTTTACCTAGAGATCTTACCAAGGAGA	549		
Db	481	CCGAGGGCGGCGATCGGGCCCTTCTGTCCACATGTTTACCTAGAGATCTTACCAAGGAGA	540		
Qy	550	AGGTATTAGACTCTCTGGAACCTGCTTCCGGGAGACCTGGTAATCCGAGAGAAGACTGCCGGG	609		
Db	541	AGGTATTAGACTCTCTGGAACCTGCTTCCGGGAGACCTGGTAATCCGAGAGAAGACTGCCGGG	600		
Qy	610	GGAAATCTGTATCCGGGTCTCTCCAGAAGCCCATAGTAGCTTTGCTGATTTTGGAC	669		
Db	601	GGAAATCTGTATCCGGGTCTCTCCAGAAGCCCATAGTAGCTTTGCTGATTTTGGAC	660		
Qy	670	GGCACTTCTGCGGACCGAGTCCGAATTCGGACTGTAGGAGCCACCCGGCTCAACACCGCGCT	729		
Db	661	GGCACTTCTGCGGACCGAGTCCGAATTCGGACTGTAGGAGCCACCCGGCTCAACMAWAGCGCT	720		
Qy	730	CCTCCGCGAGTCAATGTGTCTCTCGGTCAAGGTGGACCAAGCGGGAACGTTTGGCCCCAT	789		
Db	721	CCTCCGCGAGTCAATGTGTCTCTCGGTCAAGGTGGACCAAGCGGGAACGTTTGGCCCCAT	780		

[illegible]

	Query Match	83.8%; Score 934; DB 5; Length 958;
	Best Local Similarity	99.8%; Pred. No. 2.1e-246;
	Matches	956; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
Qy	7	CGCGGGCGGCTCGACGCAGCAGAGCGCGAGATGGCGCGAGCTTCAGCGG-CGGCG 65
Db	1	CGCGGGCGGCTCGACGCAGCAGAGCGCGAGATGGCGCGAGCTTCAGCGGNGCGG 60
Qy	66	ATCTCAGGAGTGGTCCGTGTCGGCTAAAGCAAGATTGAGACTCTGCTGCTCACCTCCA 125
Db	61	ATCTCAGGAGTGGTCCGTGTCGGCTAAAGCAAGATTGAGACTCTGCTGCTCACCTCCA 120
Qy	126	GCTCGGTTAAGGGTGCTGTGCGACTGCGGCCATTTTGTGGATGGAAACGCGGGAGCAAGT 185
Db	121	GCTCGGTTAAGGGTGCTGTGCGCATGCGGCCCATTTTGTGGATGGAAACGCGGGAGCAAGT 180
Qy	186	GATCCCCCTCTGTGTCGGGGCATGACAGCTGCTCTCTAGAGATTGCTTAATCGGAGGAAC 245
Db	181	GATCCCCCTCTGTGTCGGGGCATGACAGCTGCTCTCTAGAGATTGCTTAATCGGAGGAAC 240
Qy	246	CACCAGGAGACTCTCAAATACCAAGTTTGATGCGCTTCTATGGGAGAGAGTACTCAGCAG 305
Db	241	CACCAGGAGACTCTCAAATACCAAGTTTGATGCGCTTCTATGGGAGAGAGTACTCAGCAG 300
Qy	306	GACATCTATGTCAGGTTTCAGTGCAGCCCATCTTAAGGCATCTTCTGGAAGGCGAGAATGCC 365
Db	301	GACATCTATGTCAGGTTTCAGTGCAGCCCATCTTAAGGCATCTTCTGGAAGGCGAGAATGCC 360
Qy	366	AGTGTGCTTTCCTATGGACCCACAGAGCTGGGAAGA CGCACAAATGCTGGGCAGGCCCA 425
Db	361	AGTGTGCTTTCCTATGGACCCACAGAGCTGGGAAGA CGCACAAATGCTGGGCAGGCCCA 420
Qy	426	GAGCAACCTGGGGTGATCCCGGGGCTCTCATGGACCTCTCTGAGCTCACAAAGSGAGAG 485
Db	421	GAGCAACCTGGGGTGATCCCGGGGCTCTCATGGACCTCTCTGAGCTCACAAAGSGAGAG 480
Qy	486	GGTGCGGAGGCGGCGCATGGGCCCTTTCTGTCAACATGTCTTACCTAGAGATCTACCAG 545
Db	481	GGTGCGGAGGCGGCGCATGGGCCCTTTCTGTCAACATGTCTTACCTAGAGATCTACCAG 540
Qy	546	GAGAAGTATTAGACCTCTCTGACCCCTGCTTCGGGAGACCTGGTAAATCCGAGAGACTGC 605
Db	541	GAGAAGTATTAGACCTCTCTGACCCCTGCTTCGGGAGACCTGGTAAATCCGAGAGACTGC 600
Qy	606	CGGGGGAATATCTGATTCCGGGTCTCTCCCAGAAAGCCCATCAGTAGCTTTGCTGATTTT 665
Db	601	CGGGGGAATATCTGATTCCGGGTCTCTCCCAGAAAGCCCATCAGTAGCTTTGCTGATTTT 660
Qy	666	GAGCGCACTTCTCTGCAGCCAGTGAATAATGCGACTGTAGAGCCACCCGGCTCAACAG 725
Db	661	GAGCGCACTTCTCTGCAGCCAGTGAATAATGCGACTGTAGAGCCACCCGGCTCAACAG 720
Qy	726	CGCTCTCTCCGCGAGTCACTGCTGCTCTCTGCTCAAGGTGGACAGCGGGAACGTTTGCC 785
Db	721	CGCTCTCTCCGCGAGTCACTGCTGCTCTCTGCTCAAGGTGGACAGCGGGAACGTTTGCC 780
Qy	786	CCATTTTCGCGAGCGAGGGAACAATCTACCTGATTGACTTGGCTGGGTTCAGAGGACAC 845
Db	781	CCATTTTCGCGAGCGAGGGAACAATCTACCTGATTGACTTGGCTGGGTTCAGAGGACAC 840
Qy	846	CGCGGCACAGGCNAACAAAGGCCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTG 905
Db	841	CGCGGCACAGGCNAACAAAGGCCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTG 900
Qy	906	TTTTGCTCTGGGC-AAAGTGGTAGTGCCTGAATCAGGGCCTCCCTCGTGTACCTTAT 962
Db	901	TTTTGCTCTGGGC-AAAGTGGTAGTGCCTGAATCAGGGCCTCCCTCGTGTACCTTAT 958
	RESULT 10	
	AL559769	mRNA linear EST 02-APR-2001
	LOCUS	999 bp

```

DEFINITION  AL559769 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
              clone CS0DG005YF11 5-PRIME, mRNA sequence.
ACCESSION   AL559769
VERSION     AL559769.3 GI:46185141
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 999)
            Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
            Full-length cDNA libraries and normalization
            Unpublished (2001)
AUTHORS     On Feb 15, 2001 this sequence version replaced gi:31283900.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 7766.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?e=CS0DG005YF11&c=7766.r.
FEATURES    Location/Qualifiers
            1..999
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DG005YF11"
            /tissue_type="B CELLS (RAMOS CELL LINE)"
            /cell_line="RAMOS CELL LINE"
            /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
            /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
            with a NotI-oligo(dT) primer. Five prime end enriched,
            double-strand cDNA was digested with Not I and cloned into
            the Not I and EcoRV sites of the pCMVSPORT 6 vector.
            Library was not normalized."
ORIGIN
Query Match      82.6%; Score 921.4; DB 1; Length 999;
Best Local Similarity 96.8%; Pred. No. 6.2e-243;
Matches 913; Conservative 22; Mismatches 8; Indels 0; Gaps 0;

QY 6 GCGCGGGCGGCTCAGCAGCAGAGGCGACGCGAGATGGCGGAGCTTCAGCGCGGGG 65
DB 41 GCGCGGGCGGCTCAGCAGCAGAGGCGACGCGAGATGGCGGAGCTTCAGCGCGGGG 100
QY 66 ATCTCAGGAGCTGGTGGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCA 125
DB 101 ATCTCAGGAGCTGGTGGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCA 160
QY 126 GCTCGGTAAGGGTGGCTGTCGGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCA 185
DB 161 GCTCGGTAAGGGTGGCTGTCGGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCA 220
QY 186 GATCCCCCTGTGTCGGGGGATGACAGCTGCTCTCTAGAGATTGCTTAAGTGGAGAAC 245
DB 221 GATCCCCCTGTGTCGGGGGATGACAGCTGCTCTCTAGAGATTGCTTAAGTGGAGAAC 280
QY 246 CACAGGAGACTCTCAATACCAATGATGCTGCTCTATGGGAGAGGAGTACTCAGCAG 305
DB 281 CACAGGAGACTCTCAATACCAATGATGCTGCTCTATGGGAGAGGAGTACTCAGCAG 340
QY 306 GACATCTATGAGGTTTCAAGTCAGCCCATCTTAAGGCACTTGTCTGGAAGGCGAGATGCC 365
DB 341 GACATCTATGAGGTTTCAAGTCAGCCCATCTTAAGGCACTTGTCTGGAAGGCGAGATGCC 400
QY 366 AGTGTGCTTGTGCTATGGACCCACAGGAGCTGGGAGAGCGCACACAAATCTGGGCGGCCA 425
DB 401 AGTGTGCTTGTGCTATGGACCCACAGGAGCTGGGAGAGCGCACACAAATCTGGGCGGCCA 460

```

```

QY 426 GAGCAACCTGGGGTGATCCCGGGGCTCTCATGGACCTCTCTGAGCTTCAACAGGGAGGAG 485
DB 461 GAGCAACCTGGGGTGATCCCGGGGCTCTCATGGACCTCTCTGAGCTTCAACAGGGAGGAG 520
QY 486 GGTGCGGAGGGCGGCGCATATGGGCGGCTTCTGTCAACATGCTTTACTAGAGATCTACCCAG 545
DB 521 GGTGCGGAGGGCGGCGCATATGGGCGGCTTCTGTCAACATGCTTTACTAGAGATCTACCCAG 580
QY 546 GAGAGGATTATAGACTCTCTGAGACCTGCTTGGGAGAGCTGGTAAATCCGAGAGAGCTGC 605
DB 581 GAGAGGATTATAGACTCTCTGAGACCTGCTTGGGAGAGCTGGTAAATCCGAGAGAGCTGC 640
QY 606 CCGGGGAATATCTCTGATTTCCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTT 665
DB 641 CCGGGGAATATCTCTGATTTCCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTT 700
QY 666 GAGCGGCACTTCTCTCGACCGCATGCGAAATCGGACTGTAGGAGCCACCGGCTCAACCCAG 725
DB 701 GAGCGGCACTTCTCTCGACCGCATGCGAAATCGGACTGTAGGAGCCACCGGCTCAACCCAG 760
QY 726 CGCTCTCCCGCAGTCATGCTGCTCTCTGCTCAAGGTGGACGCGGGAAGCTTTGGCC 785
DB 761 CGCTCTCCCGCAGTCATGCTGCTCTCTGCTCAAGGTGGACGCGGGAAGCTTTGGCC 820
QY 786 CCATTTGCGCAGGAGAGGGAATCTTACCTGATTGCTTGGCTGGGTGAGAGGACCAAC 845
DB 821 CCATTTGCGCAGGAGAGGGAATCTTACCTGATTGCTTGGCTGGGTGAGAGGACCAAC 880
QY 846 CCGCGCAGGAGCAACAGGCGCTTCTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTG 905
DB 881 CCGCGCAGGAGCAACAGGCGCTTCTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTG 940
QY 906 TTTGCTCTGGGCAAGTGTAGTGGCTGAATCAAGGCGCTCC 948
DB 941 TTTGCTCTGGGCAAGTGTAGTGGCTGAATCAAGGCGCTCC 983

RESULT 11
BX376199
LOCUS      1094 bp      mRNA      linear      EST 27-APR-2004
DEFINITION Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
              cDNA clone CS0DC024YA10 5-PRIME, mRNA sequence.
ACCESSION  BX376199
VERSION    BX376199.2 GI:46616791
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 1094)
            Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
            Full-length cDNA libraries and normalization
            Unpublished (2001)
            On May 8, 2003 this sequence version replaced gi:30452476..
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            7766.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?e=CS0DC024YA10&c=7766.r.
FEATURES    Location/Qualifiers
            1..1094
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DC024YA10"

```

Query Match	82.6%	Score	921.4	DB	5	Length	1094
Best Local Similarity	99.2%	Pred. No.	6.4e-243				
Matches	963	Conservative	4	Mismatches	0	Indels	4
Qy	6	GC	CGCGGGCGCTCGACG	CAGCAGCGAGCGAGATGGCGGAGCTT	CAGCG-C	CGCG	64
Db							
Qy	12	GC	CGCGGGCGCTCGACG	CAGCAGCGCGAGAGCGGCGAGCTT	CAGCGGCGG	CGCG	71
Db							
Qy	65	GAT	CTCAGGAGCTGGT	CGCTGTCGGCTAAGCAAGATTGGAGCTACT	CGTCTGCTCCACCTCC	124	
Db							
Qy	72	GAT	CTCAGGAGCTGGT	CGCTGTCGGCTAAGCAAGATTGGAGCTACT	CGTCTGCTCCACCTCC	131	
Db							
Qy	125	AG	CTCGGTAAGGGTGGCTG	CGACTGCGGCCAATTTGTGATGGAAACAGCGGGAGCAAG	184		
Db							
Qy	132	AG	CTCGGTAAGGGTGGCTG	CGACTGCGGCCAATTTGTGATGGAAACAGCGGGAGCAAG	191		
Db							
Qy	185	TGA	TCCCCCTGTGTGCGGGGCATGGACAGCTGCTCT	CTAGAGATTGCTAACTGAGGAA	244		
Db							
Qy	192	TGA	TCCCCCTGTGTGCGGGGCATGGACAGCTGCTCT	CTAGAGATTGCTAACTGAGGAA	251		
Db							
Qy	245	CC	ACGAGAGACTCTCAAATACCAAGTTTGATGCGCTTCTAT	GGGAGAGAGGAGTACTCAGCA	304		
Db							
Qy	252	CC	ACGAGAGACTCTCAAATACCAAGTTTGATGCGCTTCTAT	GGGAGAGAGGAGTACTCAGCA	311		
Db							
Qy	305	GG	ACATCTATGCAAGTTCA	GTGAGCCCATCCTAAGCACCTTGCTG	TGGAAGGCGCAATGC	364	
Db							
Qy	312	GG	ACATCTATGCAAGTTCA	GTGAGCCCATCCTAAGCACCTTGCTG	TGGAAGGCGCAATGC	371	
Db							
Qy	365	CAG	TGCTTGCCTATGACCCACAGAGAGCTGGGAAGCAGCACAAATG	CTGGGCGAGCC	424		
Db							
Qy	372	CAG	TGCTTGCCTATGACCCACAGAGAGCTGGGAAGCAGCACAAATG	CTGGGCGAGCC	431		
Db							
Qy	425	AGA	GCAACCTGGGGTGATCCCGGGGCTCTCATGGACCTCTG	CAGCTCACAAGGGAGGA	484		
Db							
Qy	432	AGA	GCAACCTGGGGTGATCCCGGGGCTCTCATGGACCTCTG	CAGCTCACAAGGGAGGA	491		
Db							
Qy	485	GG	TGCCGAGGGCCGGCCATGGGCGCTTCTGTCAACATGTC	TACCTAGAGATCTACCA	544		
Db							
Qy	492	GG	TGCCGAGGGCCGGCCATGGGCGCTTCTGTCAACATGTC	TACCTAGAGATCTACCA	551		
Db							
Qy	545	GG	AAGGATATAGACCTCTCGACCTGCTTCGGGACCTGGTAA	TCCGAGAAGACTG	604		
Db							
Qy	552	GG	AAGGATATAGACCTCTCGACCTGCTTCGGGACCTGGTAA	TCCGAGAAGACTG	611		
Db							
Qy	605	CC	GGGGGAATATCTGATTC	CGGGTCTCTCCAGAAAGCCATCAGTAGCTTT	TGCTGATTT	664	
Db							
Qy	612	CC	GGGGGAATATCTGATTC	CGGGTCTCTCCAGAAAGCCATCAGTAGCTTT	TGCTGATTT	671	
Db							
Qy	665	TG	AGCGCATCTTC	TGCGCAGCAGTCGAAATCGAGCTG	TAGAGCCACCGGCTCAACCA	724	
Db							
Qy	672	TG	AGCGCATCTTC	TGCGCAGCAGTCGAAATCGAGCTG	TAGAGCCACCGGCTCAACCA	731	
Db							
Qy	725	GC	GCTCTCCGCGAGTATGCTGTCTCTCGTCAAGGTGAC	CACAGCGGGAACTTTGCG	784		
Db							
Qy	732	GC	GCTCTCCGCGAGTATGCTGTCTCTCGTCAAGGTGAC	CACAGCGGGAACTTTGCG	791		
Db							
Qy	785	CC	CAATTCGCGCAGAGGGGAAACTTACCTGATTC	ATGACTTGGCTGGGTGAGAGGACAA	844		
Db							
Qy	792	CC	CAATTCGCGCAGAGAGGGGAAACTTACCTGATTC	ATGACTTGGCTGGGTGAGAGGACAA	851		
Db							
Qy	845	CC	GGCGCAGAGGCAACAAAGGCGCTTCGGCT-AAA	AGAGGTGGAGCCATCAACACCTCC	903		
Db							

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12711 row: j column: 08
High quality sequence stop: 686.

FEATURES

source

1. .1019
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5723623"
/lab_host="DH10B"
/clone_lib="NIH MGC 125"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

ORIGIN

Query Match 81.5%; Score 909; DB 4; Length 1019;
Best Local Similarity 97.5%; Pred. No. 1.7e-239;
Matches 965; Conservative 0; Mismatches 21; Indels 4; Gaps 4;

```
QY 417 GCAGAGCCAGAGCAACCTGGGGTGATCCCGCGGGCTCTCATGAGACCTCTCTGAGCTCAC 476
DB 241 GCAGAGCCAGAGCAACCTGGGGTGATCCCGCGGGCTCTCATGAGACCTCTCTGAGCTCAC 300
QY 477 AGGAGGAGGGTCCGAGGCGCGGCCCATGGGCCCTTCTGTGACCAATGCTTACCTAGAG 536
DB 301 AGGAGGAGGGTCCGAGGCGCGGCCCATGGGCCCTTCTGTGACCAATGCTTACCTAGAG 360
QY 537 ATCTACAGAGAGGTATTAGACCTCTCGGAGCCCTGCTTCCGGAGACCTGGTAATCCGA 596
DB 361 ATCTACAGAGAGGTATTAGACCTCTCGGAGCCCTGCTTCCGGAGACCTGGTAATCCGA 420
QY 597 GAAGACTCGCGGGGAATATCTGATTTCCGGGTCTCTCCAGAGGCCCATCATGAGCTTT 656
DB 421 GAAGACTCGCGGGGAATATCTGATTTCCGGGTCTCTCCAGAGGCCCATCATGAGCTTT 480
QY 657 GCTGATTTGAGGGGACATCTCTGCCAGCAGTCGAAATCGCACTGTAGAGGCCACCGG 716
DB 481 GCTGATTTGAGGGGACATCTCTGCCAGCAGTCGAAATCGCACTGTAGAGGCCACCGG 540
QY 717 CTCACAGCGCTCTCCCGCAGTCATGCTGTGCTCTGCTCAAGTGGACCGGGAA 776
DB 541 CTCACAGCGCTCTCCCGCAGTCATGCTGTGCTCTGCTCAAGTGGACCGGGAA 600
QY 777 CGTTTGGCCCATTTTCGCCAGCGAGAGGGAATCTTACTGATTTGCTTGGCTGCTCA 836
DB 601 CGTTTGGCCCATTTTCGCCAGCGAGAGGGAATCTTACTGATTTGCTTGGCTGCTCA 660
QY 837 GAGGACAACCGCGGCACAGGCAACAGGGCCCTTCGGCTAAAGAGAGTGGAGCCATCAAC 896
DB 661 GAGGACAACCGCGGCACAGGCAACANAGGCCCTTCGGCTAAAGAGAGTGGAGCCATCAAC 720
QY 897 ACCTCCCTGTTGCTCTGGGCAAGTGTAGATGCGTGAATCAGGGCTCTCCTGTGTA 956
DB 721 ACCTCCCTGTTGCTCTGGGCAAGTGTAGATGCGTGAATCAGGGCTCTCCTGTGTA 780
QY 957 CTTATCGGAGCAGCAAGCTCACTCGCCATTATTCAGGAGCTCTCTGGGTGGCTCAGCCAC 1016
DB 781 CTTATCGGAGCAGCAAGCTCACTCGCCATTATTCAGGAGCTCTCTGGGTGGCTCAGCCAC 840
QY 1017 AGTATCTTTATGCAACATTCGCCCTGAGAGACGCTTCTACTAGACACAGTCTCCGCA 1076
DB 841 AGTATCTTTATGCAACATTCGCCCTGAGAGACGCTTCTACTAGACACAGTCTCCGCA 900
QY 1077 CTCACCTTTGTCGCCAGG-TCCAAGGAGTGATCAATTG 1114
DB 901 CTCACCTTTGTCGCCAGGTTCCAAGGAGTGATCAATCG 939
```

RESULT 13
BM546877
LOCUS BM546877 1019 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT_6491037 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5723623
5', mRNA sequence.
ACCESSION BM546877
VERSION BM546877.1 GI:18780186
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1019)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

This Page Blank (uspto)

Result No.	Query			ID	Description
	Score	Match	Length		
1	1115	100.0	1115	10	Adc23337 DNA encod
2	1115	100.0	1115	12	Adq60227 Human mic
3	1112.4	99.8	1538	10	Adc23339 DNA encod
4	1112.4	99.8	1538	12	Adq60229 Human mic
5	1107.4	99.3	2097	12	Adq09241 Human KNS
6	1107.4	99.3	2099	10	Adk40995 Novel hum
7	1041	93.4	2104	3	Aaf15853 Human pro
8	1038	93.1	1041	10	Adc23341 DNA encod
9	1038	93.1	1041	12	Adq60231 Human mic
10	1035.4	92.9	1464	10	Adc23343 DNA encod
11	1035.4	92.9	1464	12	Adq60233 Human mic
12	265	23.8	464	9	Ach44493 Human foe
13	233.2	20.9	531	12	ACH77408
14	230.4	20.7	232	12	ACH91108 Human gen
15	184.6	16.6	2095	11	Adm01564 Human GDN
16	184.6	16.6	3624	12	Ado44181 DNA encod
17	155.8	14.0	2374	10	Adj95073 Novel NOV
18	154.8	13.9	2375	8	Abx34596 Human mod
19	136.6	12.3	1152	6	AbA94616 Nucleotid
20	136.6	12.3	1152	8	ABBS57216 DNA encod
21	136.6	12.3	4108	6	AbA94614 Human kin

CC activity, a method of testing for ATPase activity of microtubule motor
CC proteins, methods to identify candidate agents that bind to a target
CC protein or act as a modulator of the binding characteristics or
CC biological activity of a target protein, modulators of the target
CC protein, and methods of treating cellular proliferation disorders such as
CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
CC and inflammation, for treating disorders associated with kinesin-like DNA
CC binding protein (KID) and for inhibiting KID. The sequences are used for
CC screening for modulators of motor proteins useful for treating cellular
CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
CC hypertrophy, immune disorders and inflammation, for treating disorders
CC associated with KID and for inhibiting KID and for treating autoimmune
CC diseases, arthritis, graft rejection, inflammatory bowel disease and
CC proliferation induced after medical procedures including surgery and
CC angioplasty. This sequence represents DNA encoding a human microtubule
CC motor protein of the invention. Note: The specification states that this
CC sequence encodes the protein featured as SEQ ID NO:2, but this does not
CC appear to be the case.

SQ Sequence 1115 BP; 241 A; 318 C; 333 G; 223 T; 0 U; 0 Other;

Query Match 100.0%; Score 1115; DB 12; Length 1115;
Best Local Similarity 100.0%; Pred. No. 1.6e-311; Indels 0; Gaps 0;
Matches 1115; Conservative 0; Mismatches 0;

QY 1 ATGCAGCGCGCGGCTCGACGACGAGCGAGCGCGGATGGCGGAGCTTTCAGCGG 60
DB 1 ATGCAGCGCGCGCGGCTCGACGACGAGCGAGCGCGGATGGCGGAGCTTTCAGCGG 60
QY 61 CGCGGATCTCAGAGCTGCTCGGCTAAGCAAGATTTGGAGTCTCGTCTCCAC 120
DB 61 CGCGGATCTCAGAGCTGCTCGGCTAAGCAAGATTTGGAGTCTCGTCTCCAC 120
QY 121 CTCACGCTCGGTAAGGGTGCTGCGACTCGCGGCAATTTGGATGGAACAGCGGGAG 180
DB 121 CTCACGCTCGGTAAGGGTGCTGCGACTCGCGGCAATTTGGATGGAACAGCGGGAG 180
QY 181 CAAAGTATCCCCCTGTGTCGGGGCATGGACAGCTGCTCTTAGAGATTGCTAACTGGA 240
DB 181 CAAAGTATCCCCCTGTGTCGGGGCATGGACAGCTGCTCTTAGAGATTGCTAACTGGA 240
QY 241 GGAACACACAGGAGACTTCAATACAGTTTGATGCTCTTAGAGGGAGAGGATCTC 300
DB 241 GGAACACACAGGAGACTTCAATACAGTTTGATGCTCTTAGAGGGAGAGGATCTC 300
QY 301 AGCAGACATCTATGACAGTTTCAGTCAGCCCATCTTAAGGCATTTGCTGGAAGGCACA 360
DB 301 AGCAGACATCTATGACAGTTTCAGTCAGCCCATCTTAAGGCATTTGCTGGAAGGCACA 360
QY 361 ATGCCAGTGTGCTTGCTATGGACCCACAGGAGCTGGGAAGACGACACAAATGCTGGGCA 420
DB 361 ATGCCAGTGTGCTTGCTATGGACCCACAGGAGCTGGGAAGACGACACAAATGCTGGGCA 420
QY 421 GCCCAGAGCAACTCGGGGTGATCCCGGGGCTCTCATGACCTCTCTGACGCTCACAAAGG 480
DB 421 GCCCAGAGCAACTCGGGGTGATCCCGGGGCTCTCATGACCTCTCTGACGCTCACAAAGG 480
QY 481 AGGAGGTGCGGAGGCGCGGCAATGCGCCCTTCTGTCACCATGCTTACCTAGATCT 540
DB 481 AGGAGGTGCGGAGGCGCGGCAATGCGCCCTTCTGTCACCATGCTTACCTAGATCT 540
QY 541 ACCAGAGAGGATTTAGACCTCTCTGACCTGCTTCGGGAGACCTGGTAATCCGAGAAG 600
DB 541 ACCAGAGAGGATTTAGACCTCTCTGACCTGCTTCGGGAGACCTGGTAATCCGAGAAG 600
QY 601 ACTGCCGGGGAATATCTGATTCGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTG 660
DB 601 ACTGCCGGGGAATATCTGATTCGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTG 660
QY 661 ATTTTGGGGCACTTCTCGGAGCGGAGCTCGAATCGGACTGTAGAGGCCACCCGCTCA 720
DB 661 ATTTTGGGGCACTTCTCGGAGCGGAGCTCGAATCGGACTGTAGAGGCCACCCGCTCA 720

QY 721 ACCAGCGCTCTCCCGCAGTCATGCTGTGCTCCTGGTCAAGGTGGACCGGGAAGCTT 780
DB 721 ACCAGCGCTCTCCCGCAGTCATGCTGTGCTCCTGGTCAAGGTGGACCGGGAAGCTT 780
QY 781 TGGCCCCATTTGCCAGCAGAGAGGGAATACTTACCTGATTGACTTGGCTGGGTCAAGG 840
DB 781 TGGCCCCATTTGCCAGCAGAGAGGGAATACTTACCTGATTGACTTGGCTGGGTCAAGG 840
QY 841 ACAACCGCGGCGCAGAGCAACAAGGGCTTTCGGCTAAAGAGAGCTGGAGCCATCAACACCT 900
DB 841 ACAACCGCGGCGCAGAGCAACAAGGGCTTTCGGCTAAAGAGAGCTGGAGCCATCAACACCT 900
QY 901 CCTGTTTCTCTGGGCAAGTGTAGATGCGCTGAATCAGGGCCTCCCTCGTGTACCTT 960
DB 901 CCTGTTTCTCTGGGCAAGTGTAGATGCGCTGAATCAGGGCCTCCCTCGTGTACCTT 960
QY 961 ATCGGAGACGACAGCTCACTCGCTTATTCAGAGCTCTCTGGTGGCTAGCCCAAGTA 1020
DB 961 ATCGGAGACGACAGCTCACTCGCTTATTCAGAGCTCTCTGGTGGCTAGCCCAAGTA 1020
QY 1021 TCTTATTGCCAAACATTGCCCTTGAGAGACGCTTCTACCTAGACAGCTCTCCGCACTCA 1080
DB 1021 TCTTATTGCCAAACATTGCCCTTGAGAGACGCTTCTACCTAGACAGCTCTCCGCACTCA 1080
QY 1081 ACTTTGCTGCCAGGTCACAGGAGGTGATCAATTGA 1115
DB 1081 ACTTTGCTGCCAGGTCACAGGAGGTGATCAATTGA 1115
RESULT 3
ADC23339
ID ADC23339 standard; DNA; 1538 BP.
XX
AC ADC23339;
XX
DT 18-DEC-2003 (first entry)
XX
DE DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 3).
XX
KW human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;
KW cytoskeletal; cardiant; immunomodulator; antiinflammatory; gene therapy;
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;
KW cardiac hypertrophy; immune disorder; inflammation.
OS Homo sapiens.
XX
FH Key
FT CDS
FT Location/Qualifiers
FT 1..1538
FT /product= "KID protein"
FT /transl_except= (pos: 4..5; aa: Pro)
FT /note= "This codon has an apparent 1 nucleotide deletion
FT that alters the reading frame"
XX
PN US6387644-B1.
XX
XX 14-MAY-2002.
XX
XX 28-NOV-2000; 2000US-00724224.
XX
XX 20-APR-1999; 99US-00295612.
XX 20-JUN-2000; 2000US-00597292.
XX
XX (CYTO-) CYTOKINETICS INC.
XX
XX Beraud C;
XX
XX WPI; 2003-706919/67.
XX P-PSDB; ADC23340.
XX
XX Identifying a candidate agent as modulator of function of a target
XX protein for treating cellular proliferation disorders by adding a
XX candidate agent to a mixture of the target protein that

CC activity, a method of testing for ATPase activity of microtubule motor
 CC proteins, methods to identify candidate agents that bind to a target
 CC protein or act as a modulator of the binding characteristics or
 CC biological activity of a target protein, modulators of the target
 CC protein, and methods of treating cellular proliferation disorders such as
 CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
 CC and inflammation, for treating disorders associated with kinesin-like DNA
 CC binding protein (KID) and for inhibiting KID. The sequences are used for
 CC screening for modulators of motor proteins useful for treating cellular
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
 CC hypertrophy, immune disorders and inflammation, for treating disorders
 CC associated with KID and for inhibiting KID and for treating autoimmune
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and
 CC proliferation induced after medical procedures including surgery and
 CC angioplasty. This sequence represents DNA encoding a human microtubule
 CC motor protein of the invention. Note: The specification states that this
 CC sequence encodes the protein featured as SEQ ID NO:4, but this does not
 CC appear to be the case.

XX SQ Sequence 1538 BP; 355 A; 432 C; 459 G; 292 T; 0 U; 0 Other;

Query Match 99.8%; Score 1112.4; DB 12; Length 1538;

Best Local Similarity 99.9%; Pred. No. 1.1e-310; Indels 0; Gaps 0;

Matches 1113; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCAGCGCGCGGCTCGACGACGAGCGGCGGCGGAGATGCGGCGAGCTTCAGCGG 60
 DB 1 ATGCAGCGCGCGGCTCGACGACGAGCGGCGGCGGAGATGCGGCGAGCTTCAGCGG 60
 QY 61 CGGCGATCTCAGGAGCTGCTCGCTGTCGGCTAAGCAAGATTGGAGTACTCGTCTCCAC 120
 DB 61 CGGCGATCTCAGGAGCTGCTCGCTGTCGGCTAAGCAAGATTGGAGTACTCGTCTCCAC 120
 QY 121 CTCACGCTCGCTAAGGCTGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCT 180
 DB 121 CTCACGCTCGCTAAGGCTGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCT 180
 QY 181 CAAGTGATCCCCCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTA 240
 DB 181 CAAGTGATCCCCCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTA 240
 QY 241 GGAACACAGAGAGACTCAAAATACAGATTGATGCTTCTATGCGGAGAGAGTACTC 300
 DB 241 GGAACACAGAGAGACTCAAAATACAGATTGATGCTTCTATGCGGAGAGAGTACTC 300
 QY 301 AGCAGGACATCTATCAGGCTTCAAGTGCAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA 360
 DB 301 AGCAGGACATCTATCAGGCTTCAAGTGCAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA 360
 QY 361 ATGCCAGTGTGCTTGCCTATGACCCACAGGAGCTGGGAGACGACACAAATGCTGGGCA 420
 DB 361 ATGCCAGTGTGCTTGCCTATGACCCACAGGAGCTGGGAGACGACACAAATGCTGGGCA 420
 QY 421 GCCCAGAGAACCTGGGTGATTCGCGGGCTTCTCATGACCTCTGAGCTTCAACAGGG 480
 DB 421 GCCCAGAGAACCTGGGTGATTCGCGGGCTTCTCATGACCTCTGAGCTTCAACAGGG 480
 QY 481 AGGAGGTGCGGAGGCGGCGGCTTCTGCTCACCATGCTTACCTAGATGATCT 540
 DB 481 AGGAGGTGCGGAGGCGGCGGCTTCTGCTCACCATGCTTACCTAGATGATCT 540
 QY 541 ACCAGGAGAGGTATTAGACCTTCTGACCTTCTGCGGAGACCTGGTAAATCCGAGAG 600
 DB 541 ACCAGGAGAGGTATTAGACCTTCTGACCTTCTGCGGAGACCTGGTAAATCCGAGAG 600
 QY 601 ACTGCGGGGGAATATCTGATTCGCGGTCTCTCCAGAGCCCATCAGTACTGTTGCTG 660
 DB 601 ACTGCGGGGGAATATCTGATTCGCGGTCTCTCCAGAGCCCATCAGTACTGTTGCTG 660
 QY 661 ATTTTGAGGGCACTTCTCGAGCGGCTCGAAATCGGACTGTAGAGCCACCGGCTCA 720
 DB 661 ATTTTGAGGGCACTTCTCGAGCGGCTCGAAATCGGACTGTAGAGCCACCGGCTCA 720

QY 721 ACCAGCGCTCTCCGCGAGTCAATGCTGTCTCTGCTCAAGGTGGACCGCGGAAACGTT 780
 DB 721 ACCAGCGCTCTCCGCGAGTCAATGCTGTCTCTGCTCAAGGTGGACCGCGGAAACGTT 780
 QY 781 TGGCCCCCATTTGCGCAGCGAGAGGAAACTCTACTGATTGACTTTGGTGGGTCTCAGAGG 840
 DB 781 TGGCCCCCATTTGCGCAGCGAGAGGAAACTCTACTGATTGACTTTGGTGGGTCTCAGAGG 840
 QY 841 ACAACCGGGCGCACAGCAACAGAGGCGCTTCGGCTAAAGAGAGTGGAGCCNTCAACACCT 900
 DB 841 ACAACCGGGCGCACAGCAACAGAGGCGCTTCGGCTAAAGAGAGTGGAGCCNTCAACACCT 900
 QY 901 CCTCTGTTGCTCTGCGGCAAGTGTAGATGCGCTCAATCAGGCGCTTCCCTCGGTGACTCT 960
 DB 901 CCTCTGTTGCTCTGCGGCAAGTGTAGATGCGCTCAATCAGGCGCTTCCCTCGGTGACTCT 960
 QY 961 ATCGGGACAGCAAGCTCACTGCGCTTATTCAGAGACTCTCTGCGGTGGCTCAGCCACAGTA 1020
 DB 961 ATCGGGACAGCAAGCTCACTGCGCTTATTCAGAGACTCTCTGCGGTGGCTCAGCCACAGTA 1020
 QY 1021 TCCTTATTGCGCAACATTGCGCTGAGAGACGCTTCTACCTAGACAGTCTCCGCACTCA 1080
 DB 1021 TCCTTATTGCGCAACATTGCGCTGAGAGACGCTTCTACCTAGACAGTCTCCGCACTCA 1080
 QY 1081 ACTTTGCTGCCAGGTCCAGGAGGTGATCAATTG 1114
 DB 1081 ACTTTGCTGCCAGGTCCAGGAGGTGATCAATCG 1114
 RESULT 5
 ID ADQ09241 standard; cDNA; 2097 BP.
 XX AC ADQ09241;
 XX DT 23-SEP-2004 (first entry)
 XX DE Human KNSL4 encoding cDNA SEQ ID NO:426.
 KW thanatos-associated protein; THAP; THAP responsive gene; THAP family;
 KW THAP responsive element; angiogenesis; inflammation; metastasis; cancer;
 KW apoptosis; cardiovascular disease; neurodegenerative disease; chemokine;
 KW antiangiogenic; antiinflammatory; cardiovascular; cytostatic;
 KW neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;
 KW human; gene; ss.
 XX OS Homo sapiens.
 XX PN WO2004055050-A2.
 XX PD 01-JUL-2004.
 XX PF 10-DEC-2003; 2003WO-IB006434.
 XX PR 10-DEC-2002; 2002US-0432699P.
 XX PR 03-JUL-2003; 2003US-0485027P.
 XX PA (ENDO-) ENDOCUBE SAS.
 XX PA (CNRS) CNRS CENT NAT RECH SCI.
 XX PI Girard J, Amalric F, Roussigne M, Clouaire T;
 XX WPI; 2004-525034/50.
 XX DR P-PSDB; ADQ09240.
 XX PT Modulating expression of a Thanatos (death)-Associated Protein (THAP)
 XX responsive gene for preventing or treating e.g. cancer or inflammation,
 XX comprises modulating the interaction of a THAP polypeptide with a nucleic
 XX acid.
 XX Example 47; SEQ ID NO 426; 612pp; English.
 XX The present invention describes a method for modulating the expression of

CC a thanatos (death)-associated protein (THAP) responsive gene. The method
 CC comprises modulating the interaction of a THAP-family polypeptide or its
 CC biological fragment with a nucleic acid, and so enhancing or repressing
 CC the expression of the THAP responsive gene. Also described: (1) a method
 CC of modulating the expression of a gene responsive to a THAP/chemokine
 CC complex; (2) a pharmaceutical composition comprising a THAP responsive
 CC element in a pharmaceutical carrier; (3) a transcription factor decoy
 CC comprising essentially of a THAP responsive element; (4) a cell
 CC modulating the interaction between a nucleic acid and a THAP-family
 CC polypeptide or its biological fragment, or a nucleic acid and a
 CC THAP/chemokine complex; (6) a vector packaging cell line comprising a
 CC cell comprising a viral vector which comprises a promoter operably linked
 CC to a nucleic acid encoding a THAP-family polypeptide or its biological
 CC fragment; (7) a method of constructing a cell which expresses a
 CC recombinant THAP-family polypeptide; (8) a method of ameliorating
 CC symptoms associated with a condition mediated by a THAP/chemokine complex
 CC ; (9) methods of identifying a test compound that modulates transcription
 CC at a THAP responsive element or that modulates the transport of a
 CC chemokine into the nucleus; (10) methods for reducing the symptoms
 CC associated with a condition selected from excessive or insufficient
 CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
 CC or insufficient apoptosis, cardiovascular disease and neurodegenerative
 CC diseases; symptoms associated with a condition resulting from the
 CC activity of a chemokine or a THAP-family polypeptide in an individual; or
 CC symptoms associated with transcriptional repression or activation
 CC mediated by a THAP-family polypeptide in an individual; (11) a vector
 CC comprising a THAP responsive promoter operably linked to a nucleic acid
 CC encoding a detectable product; (12) a genetically engineered cell
 CC comprising the vector described above or that expresses a THAP-family
 CC polypeptide or its biological fragment; (13) an in vitro transcription
 CC reaction comprising a nucleic acid comprising a THAP responsive promoter,
 CC ribonucleotides and an RNA polymerase; and (14) an isolated mutant THAP-
 CC family polypeptide that does not bind to a chemokine. The pharmaceutical
 CC composition has antiangiogenic, antiinflammatory, cardiovascular,
 CC cytoskeletal, neuroprotective and osteopathic activities, and can be used
 CC as a THAP and THAP synthesis modulator. The composition can be used for
 CC modulating the expression of a THAP responsive gene. Modulation is useful
 CC for reducing symptoms of conditions such as excessive or insufficient
 CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
 CC or insufficient apoptosis, cardiovascular disease or neurodegenerative
 CC diseases. The present sequence is used in the exemplification of the
 CC present invention.

XX SQ Sequence 2097 BP; 483 A; 592 C; 628 G; 394 T; 0 U; 0 Other;

Query Match 99.3%; Score 1107.4; DB 12; Length 2097;
 Best Local Similarity 99.9%; Pred. No. 3.4e-309;
 Matches 1108; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 GCCGGGGCGCTGACGACGAGCGAGCGGCGGAGATGGCGGAGCTTCAGCGGGCGG 65
 Db 28 GCCGGGGCGCTGACGACGAGCGAGCGGCGGAGATGGCGGAGCTTCAGCGGGCGG 87

Qy 66 ATCTCAGAGCTGGTGGCTGCTGCTAGCAAGATTGGAGTACTCGTGGTCCACCTCCA 125
 Db 88 ATCTCAGAGCTGGTGGCTGCTGCTAGCAAGATTGGAGTACTCGTGGTCCACCTCCA 147

Qy 126 GCTCGGTAAGGGTGGCTGTCGACTGCGGCCATTGTTGGATGGAACAGCGGGAGCAAGT 185
 Db 148 GCTCGGTAAGGGTGGCTGTCGACTGCGGCCATTGTTGGATGGAACAGCGGGAGCAAGT 207

Qy 186 GATCCCCCTGTGGGGGATGGACAGCTGCTCTAGAGATTGCTTAACCTGGAGAAC 245
 Db 208 GATCCCCCTGTGGGGGATGGACAGCTGCTCTAGAGATTGCTTAACCTGGAGAAC 267

Qy 246 CACGAGGAGACTCTCAATACAGTTTGTATGCTTCTATGGGGAGAGGAGTACTCAGCAG 305
 Db 268 CACGAGGAGACTCTCAATACAGTTTGTATGCTTCTATGGGGAGAGGAGTACTCAGCAG 327

Qy 306 GACATCTATGAGGTTCTAGTGGAGCCCATCTTAAAGGCACTTCTGGAGGCGCAATGCC 365
 Db 328 GACATCTATGAGGTTCTAGTGGAGCCCATCTTAAAGGCACTTCTGGAGGCGCAATGCC 387

Qy 366 AGTGTGCTTGGCTATGAGCCACAGAGCTGGGAGAGCCACACATGCTGGGACCCCA 425
 Db 388 AGTGTGCTTGGCTATGAGCCACAGAGCTGGGAGAGCCACACATGCTGGGACCCCA 447

Qy 426 GAGCAACCTGGGGTGTATCCCGGGGCTCTCATGGACCTCTCTGAGCTCAAGAGGAGGAG 485
 Db 448 GAGCAACCTGGGGTGTATCCCGGGGCTCTCATGGACCTCTCTGAGCTCAAGAGGAGGAG 507

Qy 486 GGTGCGAGGGCGGGCCATCTGGGCGCTTCTGTCAACATGCTTCTTACCTAGATCTTACCAG 545
 Db 508 GGTGCGAGGGCGGGCCATCTGGGCGCTTCTGTCAACATGCTTCTTACCTAGATCTTACCAG 567

Qy 546 GAGAAGGATTAGACCTCTCTGAGCCCTGCTCGGGAGAGCTGCTATATCGAGAGAGACTGC 605
 Db 568 GAGAAGGATTAGACCTCTCTGAGCCCTGCTCGGGAGAGCTGCTATATCGAGAGAGACTGC 627

Qy 606 CGGGGCAATATCTCTGATTCGGGTCTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTT 665
 Db 628 CGGGGCAATATCTCTGATTCGGGTCTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTT 687

Qy 666 GAGCGGCACTTCTCTGCGCAGCTCGAATCGGACTGTAGGAGCCACCGGCTCAACAG 725
 Db 688 GAGCGGCACTTCTCTGCGCAGCTCGAATCGGACTGTAGGAGCCACCGGCTCAACAG 747

Qy 726 CGCTCTCTCCGCGAGTCTGCTGCTCTGCTCAAGGTGACCGGCGGAACTGTTGGCC 785
 Db 748 CGCTCTCTCCGCGAGTCTGCTGCTCTGCTCAAGGTGACCGGCGGAACTGTTGGCC 807

Qy 786 CCAATTTCCGCGAGGAGGAGAACTCTACCTGATTTGACTTGGCTGGGTGAGGAGCAAC 845
 Db 808 CCAATTTCCGCGAGGAGGAGAACTCTACCTGATTTGACTTGGCTGGGTGAGGAGCAAC 867

Qy 846 CGCGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 905
 Db 868 CGCGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 927

Qy 906 TTTGTTCTGGGCAAGTGGTAGATGGCTGCTGAAATCAGGGGCTCTCTGCTGTTACCTTCGG 965
 Db 928 TTTGTTCTGGGCAAGTGGTAGATGGCTGCTGAAATCAGGGGCTCTCTGCTGTTACCTTCGG 987

Qy 966 GACAGCAAGCTCACTCGCTTATTCAGGAGCTCTCTGGGTGGGTGAGGAGGAGGAGGAGGAG 1025
 Db 988 GACAGCAAGCTCACTCGCTTATTCAGGAGCTCTCTGGGTGGGTGAGGAGGAGGAGGAGGAG 1047

Qy 1026 ATTGCCAAGCTTGGCGGCTGAGAGAGCTTCTACCTAGACAGCTCTCCGCACTCAACTTT 1085
 Db 1048 ATTGCCAAGCTTGGCGGCTGAGAGAGCTTCTACCTAGACAGCTTCTCCGCACTCAACTTT 1107

Qy 1086 GCTGCCAGGTTCCAGGAGGAGTGAATCAATTG 1114
 Db 1108 GCTGCCAGGTTCCAGGAGGAGTGAATCAATCG 1136

RESULT 6
 ADK40995
 ID ADK40995 standard; DNA; 2099 BP.
 XX
 AC ADK40995;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Novel human kinase gene #15.
 KW cytosolic; immunomodulator; cardiant; neuroprotective; nootropic;
 KW antiparkinsonian; virucide; antibacterial; fungicide; ophthalmological;
 KW analgesic; hypotensive; immunosuppressive; kinase inhibitor; kinase;
 KW cancer; peripheral nervous system; central nervous system;
 KW Alzheimer's disease; Parkinson's disease; multiple sclerosis;
 KW amyotrophic lateral sclerosis; viral infection; prion infection;
 KW ocular disease; migraine; pain; sexual dysfunction; mood disorder;
 KW attention disorder; cognition disorder; hypotension; hypertension;
 KW psychotic disorder; neurological disorder; dyskinesia;

gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
wound; infectious disease; ss.
Homo sapiens.
W0200055174-A1.
21-SEP-2000.
08-MAR-2000; 2000WO-US005988.
12-MAR-1999; 99US-0124270P.
(HUMA-) HUMAN GENOME SCI INC.
(ROSE/) ROSEN C A.
Rosen CA, Ruben SM;
WPI; 2000-587513/55.
P-PSDB; AAB56650.
Prostate cancer associated gene sequences, referred to as prostate cancer
antigens, useful for treatment, prevention, and diagnosis of disorders
such as prostate cancer.
Claim 1; Page 805-806; 2338pp; English.
AAF15566 to AAF16505 encode the human prostate cancer associated
proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
The prostate cancer antigens can have neuroprotective, cytoskeletal,
cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
nephrotropic, antiinfective, gynaecological and antibacterial activities,
and can be used in gene therapy. The prostate cancer antigen
polynucleotides may be used for detection of prostate cancer, chromosome
identification, as chromosome markers, and for numerous other diagnostic
or research purposes. The prostate cancer antigens may be used to treat
disorders such as neural, immune, muscular, reproductive,
gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
AAF16515 to AAF16516 represent sequences used in the exemplification of the present
invention
Sequence 2104 BP; 517 A; 580 C; 597 G; 400 T; 0 U; 10 Other;
Query Match 93.4%; Score 1041; DB 3; Length 2104;
Best Local Similarity 99.7%; Pred. No. 5.2e-290;
Matches 1041; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
71 AGGAGCTGGTGGCTGCTGGCTAAGCAAGATTGGAGTACTGCTGCTCCACCTCCAGCTCG 130
60 AGGAGCTGGTGGCTGCTGGCTAAGCAAGATTGGAGTACTGCTGCTCCACCTCCAGCTCG 119
131 CGTAAGGGTGGCTGCTGGCTGCTGGCTGCTGGCTGCTGGCTGCTGGCTGCTGGCTGCT 190
120 CGTAAGGGTGGCTGCTGGCTGCTGGCTGCTGGCTGCTGGCTGCTGGCTGCTGGCTGCT 179
191 CCCCTGCTGGGGGCGATGGACAGCTGCTCTAGAGATTGCTAAGTGGAGGACACCA 250
180 CCCCTGCTGGGGGCGATGGACAGCTGCTCTAGAGATTGCTAAGTGGAGGACACCA 239
251 GGAGACTCTCAATACCAATGTTGATGCTTCTATGGGGAGAGAGTACTCAGCAGGACAT 310
240 GGAGACTCTCAATACCAATGTTGATGCTTCTATGGGGAGAGAGTACTCAGCAGGACAT 299
311 CTATGCAAGTTTCAAGTGCACCCATCTTAAAGCACTTGTGGAAGGCGGAGAAATGCAAGTGT 370
300 CTATGCAAGTTTCAAGTGCACCCATCTTAAAGCACTTGTGGAAGGCGGAGAAATGCAAGTGT 359
371 GCTTGCCTATGCAAGTGCACCCATCTTAAAGCACTTGTGGAAGGCGGAGAAATGCAAGTGT 430
360 GCTTGCCTATGCAAGTGCACCCATCTTAAAGCACTTGTGGAAGGCGGAGAAATGCAAGTGT 419
431 ACCTGGGGTGTATCCCGCGGGCTCTCATGGACCTCTCTGAGCTCACAAGGGAGAGGGTGC 490

Db 420 ACCTGGGGTGTATCCCGCGGGCTCTCATGGACCTCTCTGAGCTCACAAGGGAGAGGGTGC 479
Qy 491 CGAGGGCGGGCCATGGGGCCCTTCTGTCAACCATGTCTTACCTAGAGATCTTACAGAGAA 550
Db 480 CGAGGGCGGGCCATGGGGCCCTTCTGTCAACCATGTCTTACCTAGAGATCTTACAGAGAA 539
Qy 551 GGTATTAGACCTCTCTGAGCCCTGCTTCCGGAGACCTGGTAATCCGAGAGACTGCGCGGG 610
Db 540 GGTATTAGACCTCTCTGAGCCCTGCTTCCGGAGACCTGGTAATCCGAGAGACTGCGCGGG 599
Qy 611 GAATATCTGATTCGGGGTCTCTCCAGAAAGCCCATCAGTAGTCTTCTGATTTGAGCG 670
Db 600 GAATATCTGATTCGGGGTCTCTCCAGAAAGCCCATCAGTAGTCTTCTGATTTGAGCG 659
Qy 671 GCACTTCTGCGAGCCAGTCCGAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTC 730
Db 660 GCACTTCTGCGAGCCAGTCCGAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTC 719
Qy 731 CTCCCGCAGTCATGCTGTGCTCTGCTCAAGGTGGACACGCGGGAACGTTTGGCCCCATT 790
Db 720 CTCCCGCAGTCATGCTGTGCTCTGCTCAAGGTGGACACGCGGGAACGTTTGGCCCCATT 779
Qy 791 TCGCCAGCGAGAGGAAACTCTACCTGATTTGACTTGGCTGGGTTCAGAGCAACCGGCG 850
Db 780 TCGCCAGCGAGAGGAAACTCTACCTGATTTGACTTGGCTGGGTTCAGAGCAACCGGCG 839
Qy 851 CACAGGCAACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGT 910
Db 840 CACAGGCAACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGT 899
Qy 911 CTTGGGCAAAAGTGTAGTCCGCTGAATCAGGGCTCCCTCGTGTACTTATCGGACAG 970
Db 900 CTTGGGCAAAAGTGTAGTCCGCTGAATCAGGGCTCCCTCGTGTACTTATCGGACAG 959
Qy 971 CAAGCTCACTCGCTATTGTCAGAGCTCTCTGGGTGGCTCAGCCACAGTATCTTATTGC 1030
Db 960 CAAGCTCACTCGCTATTGTCAGAGCTCTCTGGGTGGCTCAGCCACAGTATCTTATTGC 1019
Qy 1031 CAACATGCGCCCTGAGAGAGCTCTTACCTAGACAGCTCTCCGCACTCAACTTTGCTGC 1090
Db 1020 CAACATGCGCCCTGAGAGAGCTCTTACCTAGACAGCTCTCCGCACTCAACTTTGCTGC 1079
Qy 1091 CAGGTCCAAGGAGGTGATCAATTG 1114
Db 1080 CAGGTCCAAGGAGGTGATCAATCG 1103
RESULT 8
ADC23341
ID ADC23341 standard; DNA; 1041 BP.
XX AC ADC23341;
XX AC ADC23341;
XX DT 18-DEC-2003 (first entry)
XX DE DNA encoding the human kinesin-like DNA binding protein (KID) (seqID 5).
XX KW human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;
KW cytoskeletal; cardiant; immunomodulator; antiinflammatory; gene therapy;
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;
KW cardiac hypertrophy; immune disorder; inflammation.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
CDS 1..1041
FT /*tag= a
FT /product= "KID protein"
XX FT US6387644-B1.
XX EN 14-MAY-2002.
XX PD

XX 28-NOV-2000; 2000US-00724224.
 XX 20-APR-1999; 99US-00295612.
 PR 20-JUN-2000; 2000US-00597292.
 XX (CYTO-) CYTOKINETICS INC.
 PA Beraud C;
 XX WPI: 2003-706919/67.
 XX P-PSDB; ADC23342.
 PT Identifying a candidate agent as modulator of function of a target
 PT protein for treating cellular proliferation disorders by adding a
 PT candidate agent to a mixture of the target protein that
 PT directly/indirectly produces ADP or phosphate.
 XX Disclosure; SEQ ID NO 5; 26pp; English.
 XX This invention relates to a novel method for high throughput screening
 CC systems used to identify compounds for the treatment of cellular
 CC proliferation disorders. Specifically, it refers to candidate agents that
 CC are capable of modulating the activity of target proteins having motor
 CC domains, such that the target protein directly or indirectly produces ADP
 CC or phosphate. Furthermore, this activity can be determined using
 CC fluorescence or absorbance readouts. The present invention describes a
 CC method that identifies modulators of the target protein, which is a
 CC kinesin-like DNA binding protein (known as KID) as cytoskeletal, cardinals,
 CC immunomodulators and antiinflammatory. Accordingly, through gene
 CC therapy, they can be used for the treatment of cancer, hyperplasias,
 CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This
 CC polynucleotide sequence is human KID DNA (SeqID 5) encoding a full length
 CC KID enzyme of the invention.
 XX
 SQ Sequence 1041 BP; 228 A; 295 C; 302 G; 216 T; 0 U; 0 Other;
 Query Match 93.1%; Score 1038; DB 10; Length 1041;
 Best Local Similarity 100.0%; Pred. No. 2.9e-289;
 Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 GGTGCTGTGCGCTAAGCAAGATTGGAGTACTGTGCTGCTCACTCCAGCTCGGTAAAGG 137
 DB 4 GGTGCTGTGCGCTAAGCAAGATTGGAGTACTGTGCTGCTCACTCCAGCTCGGTAAAGG 63
 QY 138 GTGGCTGTGCGCTGCGGCAATTTGGATGGAACAGCGGGAGCAAGTATCCCCCTGT 197
 DB 64 GTGGCTGTGCGCTGCGGCAATTTGGATGGAACAGCGGGAGCAAGTATCCCCCTGT 123
 QY 198 GTGCGGGCATGACAGCTGCTCTTAGAGATTGCTTAACCTGGAGGAACCAACAGGAGACT 257
 DB 124 GTGCGGGCATGACAGCTGCTCTTAGAGATTGCTTAACCTGGAGGAACCAACAGGAGACT 183
 QY 258 CTCAAATACAGTTTGTATGCTTTATGGGGAGAGTACTCAGCAGACATCTATGCA 317
 DB 184 CTCAAATACAGTTTGTATGCTTTCTATGGGGAGAGTACTCAGCAGACATCTATGCA 243
 QY 318 GGTTCAGTGCAGCCATCTCCTAGGCACTTGTGGAGGCAATGCTGCTGCTTGGC 377
 DB 244 GGTTCAGTGCAGCCATCTCCTAGGCACTTGTGGAGGCAATGCTGCTGCTTGGC 303
 QY 378 TATGACCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGCCACAGAGCAACTGGG 437
 DB 304 TATGACCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGCCACAGAGCAACTGGG 363
 QY 438 GTGATCCCGCGGCTCTCATAGACCTCTGAGCTCACAAGGGAGAGGGTCCGAGGCG 497
 DB 364 GTGATCCCGCGGCTCTCATAGACCTCTGAGCTCACAAGGGAGAGGGTCCGAGGCG 423
 QY 498 CGGCCATGGGCCCTTCTGTACCATCTTACCTAGAGATCTACAGGAGAGGTATTA 557
 DB 424 CGGCCATGGGCCCTTCTGTACCATCTTACCTAGAGATCTACAGGAGAGGTATTA 483

QY 558 GACCTCTGAGCCCTGCTTCCGGAGACCTGGTAATCCGAGAAGACTGCCGGGGAATATC 617
 DB 484 GACCTCTGAGCCCTGCTTCCGGAGACCTGGTAATCCGAGAAGACTGCCGGGGAATATC 543
 QY 618 CTGATTCCGGGCTCTCTCCAGAACGCCATCAGTAGCTTTGCTGATTTTGTAGCGGCACTTC 677
 DB 544 CTGATTCCGGGCTCTCTCCAGAACGCCATCAGTAGCTTTGCTGATTTTGTAGCGGCACTTC 603
 QY 678 CTGCGAGCCAGTCCGAATCCGAGCTGTAGGAGCCACCCGGCTCAACAGGCTCTCTCCCGC 737
 DB 604 CTGCGAGCCAGTCCGAATCCGAGCTGTAGGAGCCACCCGGCTCAACAGGCTCTCTCCCGC 663
 QY 738 AGTCATGCTGTGCTCTCTGCTCAAGGTGGACCGGGAAACGTTTGGCCCCCATTTCCGCCAG 797
 DB 664 AGTCATGCTGTGCTCTCTGCTCAAGGTGGACCGGGAAACGTTTGGCCCCCATTTCCGCCAG 723
 QY 798 CGAGAGGGAAACTCTTACCTGATTGACTTTGGCTGGGTGAGAGCAAAACCGGCGCACAGGC 857
 DB 724 CGAGAGGGAAACTCTTACCTGATTGACTTTGGCTGGGTGAGAGCAAAACCGGCGCACAGGC 783
 QY 858 AACAGGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGC 917
 DB 784 AACAGGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGC 843
 QY 918 AAGTGGTAGATGCGCTGAATCAGGGCCCTCCCTCGGTGTAACCTTATCGGACAGCAAGCTC 977
 DB 844 AAGTGGTAGATGCGCTGAATCAGGGCCCTCCCTCGGTGTAACCTTATCGGACAGCAAGCTC 903
 QY 978 ACTCGCTTATTCAGAGCTCTCTGCGTGGCTCAGGCCACAGTATCTTATTCGCAACATT 1037
 DB 904 ACTCGCTTATTCAGAGCTCTCTGCGTGGCTCAGGCCACAGTATCTTATTCGCAACATT 963
 QY 1038 GCCCTGTGAGAGCGCTTCTACCTAGACACAGTCTCGGCACTCAACTTGTGCTGCGAGTCC 1097
 DB 964 GCCCTGTGAGAGCGCTTCTACCTAGACACAGTCTCGGCACTCAACTTGTGCTGCGAGTCC 1023
 QY 1098 AAGGAGGTGATCAATTGA 1115
 DB 1024 AAGGAGGTGATCAATTGA 1041

RESULT 9
 ADQ60231
 ID ADQ60231 standard; DNA; 1041 BP.
 XX AC ADQ60231;
 XX DT 23-SEP-2004 (first entry)
 XX Human microtubule motor protein DNA #3.
 DE Human; microtubule motor protein; gene; ds;
 KW cellular proliferation disorder; cancer; hyperplasia; restenosis;
 KW cardiac hypertrophy; immune disorder; inflammation;
 KW kinesin-like DNA binding protein; KID; autoimmune disease; arthritis;
 KW graft rejection; inflammatory bowel disease; angioplasty.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT 1..1041
 CDS /*tag= a
 /product= "Human microtubule motor protein #3"
 PN US6762043-B1.
 XX 13-JUL-2004.
 XX 06-MAR-2002; 2002US-00093317.
 XX 20-APR-1999; 99US-00295612.
 PR 20-JUN-2000; 2000US-00597292.
 PR 28-NOV-2000; 2000US-00724224.

Fri Nov 12 12:26:08 2004

(CYTO-) CYTOKINETICS INC.

Beraud C;

WPI; 2004-532491/51.
P-PSDB; ADQ60232.

P-PSDB; ADQ60232.

New isolated microtubule motor protein, useful for screening modulators for treating cellular proliferation disorders such as cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders and inflammation.

Disclosure: SEO ID NO 5: 26pp; English.

The invention relates to human microtubule motor proteins and the nucleic acids encoding them. The invention also relates to a method of screening for modulators of a motor protein which has microtubule stimulated ATPase activity, a method of testing for ATPase activity of microtubule motor proteins, methods to identify candidate agents that bind to a target protein or act as a modulator of the binding characteristics or biological activity of a target protein, modulators of the target protein, and methods of treating cellular proliferation disorders such as cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders and inflammation, for treating disorders associated with kinesin-like DNA binding protein (KID) and for inhibiting KID. The sequences are used for screening for modulators of motor proteins useful for treating cellular proliferation disorders such as cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders and inflammation, for treating disorders associated with KID and for inhibiting KID and for treating autoimmune diseases, arthritis, graft rejection, inflammatory bowel disease and proliferation induced after medical procedures including surgery and angioplasty. This sequence represents DNA encoding a human microtubule motor protein of the invention.

Sequence 1041 BP: 228 A: 295 C: 302 G: 216 T: 0 U: 0 Other:

Query Match 93.1%; Score 1038; DB 12; Length 1041;

Best Local Similarity 100.0%; Pred. No. 2.9e-289;

Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	78	GGTGCCTGTCGGCTAAGCAAGNATGGAGCTACTCGTGGTCCACCTCCAGCTCGCGTAAGG	13
Db	4	GGTGCCTGTCGGCTAAGCAAGNATGGAGCTACTCGTGGTCCACCTCCAGCTCGCGTAAGG	63
Qy	138	GTGGCTGTGCGACTGCGGGCCATTTGTGGATGGAACAGCGGAGCAAGTGATCCCCCCTGT	197
Db	64	GTGGCTGTGCGACTGCGGGCCATTTGTGGATGGAACAGCGGAGCAAGTGATCCCCCCTGT	123
Qy	198	GTGCGGGGCATGGAACAGCTGCTCTCTAGAGATTGCTTAATCTGGAGGAACCAACGAGGAGACT	257
Db	124	GTGCGGGGCATGGAACAGCTGCTCTCTAGAGATTGCTTAATCTGGAGGAACCAACGAGGAGACT	183
Qy	258	CTCAAAATACCAATGTTGATGCCCTTCTATGGGGAGGAGTACTCAGCAGGAGCATCTATGCA	317
Db	184	CTCAAAATACCAATGTTGATGCCCTTCTATGGGGAGGAGTACTCAGCAGGAGCATCTATGCA	243
Qy	318	GGTTCAGTGCAGGCCCATCTCAAGGCACATTGCTGGAAGGGCAGAAATGCCAGTGTGCTTGCC	377
Db	244	GGTTCAGTGCAGGCCCATCTCAAGGCACATTGCTGGAAGGGCAGAAATGCCAGTGTGCTTGCC	303
Qy	378	TATGGACCCACAGGAGCTGGGAAGAGCCACACAATGCTGGGCAGGCCACGACCACTTGGG	437
Db	304	TATGGACCCACAGGAGCTGGGAAGAGCCACACAATGCTGGGCAGGCCACGACCACTTGGG	363
Qy	438	GTGATCCCGGGGCTCTCATGGAACCTCTGTCAGCTCACAAAGGAGGAGGGTGCCGAGGGC	497
Db	364	GTGATCCCGGGGCTCTCATGGAACCTCTGTCAGCTCACAAAGGAGGAGGGTGCCGAGGGC	423
Qy	498	CGGCCATGGGCCCTTTCTGTCAACCATGCTTTACCTTAGAGATCTACCGAGGAGAGGTATTA	557
Db	424	CGGCCATGGGCCCTTTCTGTCAACCATGCTTTACCTTAGAGATCTACCGAGGAGAGGTATTA	483

QY	558	GA	CTCTCTCGGAGCCCTGCTTCCGGGAGACCTGGTAATCCGAGAGACTGCGGGGGGAATATC	617			
DB	484	GA	CTCTCTCGGAGCCCTGCTTCCGGGAGACCTGGTAATCCGAGAGACTGCGGGGGGAATATC	543			
QY	618	CT	GATTTCCGGGCTCTCTCCAGAGCCCATCAGTAGTCTTTGCTGATTTTGTAGCGGCACTTC	677			
DB	544	CT	GATTTCCGGGCTCTCTCCAGAGCCCATCAGTAGTCTTTGCTGATTTTGTAGCGGCACTTC	603			
QY	678	CT	GCCAGCCAGTTCGAAATCGGACTGTAGGAGCCACCGGCTCAACAGAGCGCTCCTCCCGC	737			
DB	604	CT	GCCAGCCAGTTCGAAATCGGACTGTAGGAGCCACCGGCTCAACAGAGCGCTCCTCCCGC	663			
QY	738	AG	TCATGCTCTGCTCTCTGCTCAAGGTGGACACGGGGAACGTTTGGCCCCCATTTGGCCAG	797			
DB	664	AG	TCATGCTCTGCTCTCTGCTCAAGGTGGACACGGGGAACGTTTGGCCCCCATTTGGCCAG	723			
QY	798	CG	AGAGGGAAAACTCTACCTGATTTGACTTTGGCTGGGTTCAGAGACAAACCGGCGCACAGGC	857			
DB	724	CG	AGAGGGAAAACTCTACCTGATTTGACTTTGGCTGGGTTCAGAGACAAACCGGCGCACAGGC	783			
QY	858	AA	CAAGGGCCCTCGGCTTAAAGAGGTGGAGCCATCAACACCTCCCTGTTGTCTCTGGGC	917			
DB	784	AA	CAAGGGCCCTCGGCTTAAAGAGGTGGAGCCATCAACACCTCCCTGTTGTCTCTGGGC	843			
QY	918	AA	AGTGTAGATGCGCTGTAATCAGGGGCTCCCTCGTGTACTTATTCGGGACAGCAAGCTC	977			
DB	844	AA	AGTGTAGATGCGCTGTAATCAGGGGCTCCCTCGTGTACTTATTCGGGACAGCAAGCTC	903			
QY	978	ACT	CGCCTATTCAGGAGCTCTCTGGGTGGCTCAGCCACAGTATCCTTATTGCCAACATT	1037			
DB	904	ACT	CGCCTATTCAGGAGCTCTCTGGGTGGCTCAGCCACAGTATCCTTATTGCCAACATT	963			
QY	1038	GCC	CTCAGAGACGCTTCTACCTAGACACAGTCTCTCCGACTCAACTTTGCTGCGAGTCC	1097			
DB	964	GCC	CTCAGAGACGCTTCTACCTAGACACAGTCTCTCCGACTCAACTTTGCTGCGAGTCC	1023			
QY	1098	AAG	GAGGTGATCAATTGA	1115			
DB	1024	AAG	GAGGTGATCAATTGA	1041			
RESULT 10							
ADC23343							
ID	ADC23343 standard; DNA; 1464 BP.						
XX							
AC	ADC23343;						
XX							
DT	18-DEC-2003 (first entry)						
XX							
DE	DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 7).						
XX							
KW	human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;						
KW	cytostatic; cardiant; immunomodulator; antiinflammatory; gene therapy;						
KW	cancer; hyperplasia; restenosis; cellular proliferation disorder;						
KW	cardiac hypertrophy; immune disorder; inflammation.						
XX							
OS	Homo sapiens.						
XX							
FH	Key	Location/Qualifiers					
FT	CDS	1..1464					
FT		/*tag= a					
FT		/product= "KID protein"					
XX							
PN	US6387644-B1.						
XX							
PD	14-MAY-2002.						
XX							
PP	28-NOV-2000; 2000US-00724224.						
XX							
PR	20-APR-1999; 99US-00295612.						
XX							
PR	20-JUN-2000; 2000US-00597292.						
XX							
PA	(CYTO-) CYTOKINETICS INC.						

XX New isolated microtubule motor protein, useful for screening modulators
PT for treating cellular proliferation disorders such as cancer,
PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
PT inflammation.
PS
XX Disclosure; SEQ ID NO 7; 26pp; English.
XX The invention relates to human microtubule motor proteins and the nucleic
CC acids encoding them. The invention also relates to a method of screening
CC for modulators of a motor protein which has microtubule stimulated ATPase
CC activity, a method of testing for ATPase activity of microtubule motor
CC proteins, methods to identify candidate agents that bind to a target
CC protein or act as a modulator of the binding characteristics or
CC biological activity of a target protein, modulators of the target
CC protein, and methods of treating cellular proliferation disorders such as
CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
CC and inflammation, for treating disorders associated with kinesin-like DNA
CC binding protein (KID) and for inhibiting KID. The sequences are used for
CC screening for modulators of motor proteins useful for treating cellular
CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
CC hypertrophy, immune disorders and inflammation, for treating disorders
CC associated with KID and for inhibiting KID and for treating autoimmune
CC diseases, arthritis, graft rejection, inflammatory bowel disease and
CC proliferation induced after medical procedures including surgery and
CC angioplasty. This sequence represents DNA encoding a human microtubule
XX motor protein of the invention.
XX
XX Sequence 1464 BP; 342 A; 409 C; 428 G; 285 T; 0 U; 0 Other;

Query Match 92.9%; Score 1035.4; DB 12; Length 1464;
Best Local Similarity 99.9%; Pred. No. 1.9e-288;
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 78 GGTCTGCTGGCTAAGCAAGATTGGAGCTACTCTGCTGCTCCACCTCCAGCTCGCGTAAGG 137
DB |||||||
QY 138 GTGGCTGTGCGACTCGGCCCAATTTGTGATGGAAACAGCGGAGCAAGTGATCCCCCTGT 197
DB |||||||
QY 198 GTGCGGGCATGGACAGCTGCTCTCTAGAGATTGCTTAAGTGGAGCAACACAGAGACT 257
DB |||||||
QY 124 GTGCGGGCATGGACAGCTGCTCTCTAGAGATTGCTTAAGTGGAGCAACACAGAGACT 183
DB |||||||
QY 258 CTCAATACCACTTTGATGCTCTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 317
DB |||||||
QY 318 GTTTCAGTGCAGCCCATCTAAGGCACTTGTGGAAGGCGAGATGCCAGTGTGCTTGC 377
DB |||||||
QY 244 GGTTCAGTGCAGCCCATCTAAGGCACTTGTGGAAGGCGAGATGCCAGTGTGCTTGC 303
DB |||||||
QY 378 TATGACCCACAGGAGCTGGAGAGCGACACAATGCTGGGAGCCACAGAGCAACTGGG 437
DB |||||||
QY 304 TATGACCCACAGGAGCTGGAGAGCGACACAATGCTGGGAGCCACAGAGCAACTGGG 363
DB |||||||
QY 438 GTGATCCCGGGCTCTCATGACCTCTGAGCTTCAAGGGAGAGGGTCCGAGGCG 497
DB |||||||
QY 364 GTGATCCCGGGCTCTCATGACCTCTGAGCTTCAAGGGAGAGGGTCCGAGGCG 423
DB |||||||
QY 498 CGGCCATGGGCCCTTCTCTGCTCCATGCTTACCTAGAGATCTACAGGAGAGGTATTA 557
DB |||||||
QY 424 CGGCCATGGGCCCTTCTCTGCTCCATGCTTACCTAGAGATCTACAGGAGAGGTATTA 483
DB |||||||
QY 558 GACCTCTGGAACCTCTCTGGGAGACCTGGTAATCCGAGAGACTGCGGGGGAATATC 617
DB |||||||
QY 484 GACCTCTGGAACCTCTCTGGGAGACCTGGTAATCCGAGAGACTGCGGGGGAATATC 543
DB |||||||
QY 618 CTGATTCGGGGTCTCTCCAGAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCATTC 677
DB |||||||
QY 544 CTGATTCGGGGTCTCTCCAGAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCATTC 603
DB |||||||

QY 678 CTGCCAGCCAGTCGAAATCGAGCTGTAGGAGCCACCCGGCTCAACAGCGCTCTCTCCCGC 737
DB |||||||
QY 604 CTGCCAGCCAGTCGAAATCGAGCTGTAGGAGCCACCCGGCTCAACAGCGCTCTCTCCCGC 663
DB |||||||
QY 738 AGTCATGCTGTGCTCTCTGCTCAAGGTGACACAGCGGGAAACGTTTGGGCCCCCATTTGCCAG 797
DB |||||||
QY 664 AGTCATGCTGTGCTCTCTGCTCAAGGTGACACAGCGGGAAACGTTTGGGCCCCCATTTGCCAG 723
DB |||||||
QY 798 CGAGAGGGAACACTCTACTGATTGCTGCTGGTTCAGAGGACAAACCGGGGCAACAGC 857
DB |||||||
QY 724 CGAGAGGGAACACTCTACTGATTGCTGCTGGTTCAGAGGACAAACCGGGGCAACAGC 783
DB |||||||
QY 858 AACAGAGGCTTCGGCTTAAGAGAGTGGAGCCATCAACACCTCTGTTGCTCTGGGC 917
DB |||||||
QY 784 AACAGAGGCTTCGGCTTAAGAGAGTGGAGCCATCAACACCTCTGTTGCTCTGGGC 843
DB |||||||
QY 918 AAAGTGTAGATGCTGCTGAATCAGGGCTCTCTGCTGATCTTATCGGAGACAGAGCTC 977
DB |||||||
QY 844 AAAGTGTAGATGCTGCTGAATCAGGGCTCTCTGCTGATCTTATCGGAGACAGAGCTC 903
DB |||||||
QY 978 ACTGCGCTTATTCAGAGACTCTCTGCTGCTGCTCAGGCCACAGTATCTTATTCGCAACT 1037
DB |||||||
QY 904 ACTGCGCTTATTCAGAGACTCTCTGCTGCTGCTCAGGCCACAGTATCTTATTCGCAACT 963
DB |||||||
QY 1038 GCGCTGAGAGAGCTTCTTACCTAGACAGCTCTCGGCACTCAACTTTGCTGCGAGTCC 1097
DB |||||||
QY 964 GCGCTGAGAGAGCTTCTTACCTAGACAGCTCTCGGCACTCAACTTTGCTGCGAGTCC 1023
DB |||||||
QY 1098 AAGGAGGTGATCAATTG 1114
DB |||||||
QY 1024 AAGGAGGTGATCAATCG 1040
DB |||||||
RESULT 12
ACH4493
ID ACH44493 standard; cDNA; 464 BP.
XX
XX ACH44493;
XX
XX 13-OCT-2003 (first entry)
DT Human foetal brain cDNA #5218.
DE
XX
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
XX Homo sapiens.
OS
XX
XX US2003073623-A1.
XX
XX 17-APR-2003.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
PI
XX
XX WPI; 2003-615964/58.
DR
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 31705; 44pp; English.
PS

XX The invention relates to an isolated polynucleotide comprising any one of
 CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
 CC determined by the technique of SBH (sequencing by hybridisation). Also
 CC included is a purified polypeptide comprising a sequence corresponding to
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences
 CC are useful in diagnostics as expressed sequence tags (EST) for
 CC identifying expressed genes or for physical mapping of the human genome,
 CC in forensics, in assessing biodiversity, or in identifying mutations
 CC responsible for genetic disorders and other traits. The nucleotide
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,
 CC for chromosome and gene mapping, in the recombinant production of
 CC protein, or in generating antisense DNA or RNA. The purified polypeptide
 CC is useful for generating antibodies specific for it. The present sequence
 CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030073623
 XX
 SQ Sequence 464 BP; 96 A; 136 C; 131 G; 96 T; 0 U; 5 Other;

Query Match 23.8%; Score 265; DB 9; Length 464;
 Best Local Similarity 77.6%; Pred. No. 4.9e-66;
 Matches 333; Conservative 0; Mismatches 95; Indels 1; Gaps 1;
 QY 200 GCGGGCATGGACAGCTGCTCTAGAGATTGCTAACTGGAGAACCCACGAGACTCT 259
 Db 37 GCTCGAATTCATATCGCGTGGATAGAGTGAACGCCGACCAACCCGATATCT 96
 QY 260 CAATACACATTTGATGCTTCTATGGGAGAGAGTACTACGAGGACATCTATGCAGG 319
 Db 97 CCTTACCTGTGCTTGGCTTTGACAGCCATAGAGGAGCTGCGCTGATGATCTTCC 156
 QY 320 TTCAGTCAGCCCATCTTAAGGCACTTCTGAGGCGGAGATGCCAGTGTGCTTGCCTA 379
 Db 157 TTGACCGTGGACCGTGCCTACTGCCACTGCTCTTCTGTAATAAGCGAGGAGCTGCGCTA 216
 QY 380 TGGACCCACAGAGCTGGGAGACGACACAAATGCTGGGCGAGCCGACGACCTGGGGT 439
 Db 217 CGGTGCTGAGGAGCTGGCATTAC-ATCTGATGCTTGGCAGCCGACGACCTGGGGT 275
 QY 440 GATCCCGGGGCTCTCATGGACCTCTGAGCTCAAGAGGAGAGGTGGCGAGGGCG 499
 Db 276 GATCCCGGGGCTCTCATGGACCTCTGAGCTCAAGAGGAGAGGTGGCGAGGGCG 335
 QY 500 GCCATGGGCTTCTGTCACCATGCTTACCTAGATCTACGAGGAGAGGTATTAGA 559
 Db 336 GCCATGGGCTTCTGTCACCATGCTTACCTAGATCTACGAGGAGAGGTATTAGA 395
 QY 560 CCTCTGGACCTCTTGGGAGACCTGCTAATCCGAGAGACTGCCGGGGGAATATCT 619
 Db 396 CCTCTGGACCTCTTGGGAGACCTTGGGAGAGACTTGGGAGAGACTGCCGGGGGAATATCT 455
 QY 620 GATTCGGG 628
 Db 456 GATTCGGG 464

RESULT 13
 ACH77408
 ID ACH77408 standard; DNA; 531 BP.
 XX ACH77408;
 AC ACH77408;
 XX ACH77408;
 XX ACH77408;
 DT 29-JUL-2004 (first entry)
 XX Human genome derived single exon probe #10603.
 DE Human; probe: ss; gene expression; single exon probe; microarray;
 KW alternative splicing event; genomic alteration.
 XX Homo sapiens.
 OS
 XX

PN US2003194704-A1.
 XX 16-OCT-2003.
 PD 03-APR-2002; 2002US-00029386.
 XX 03-APR-2002; 2002US-00029386.
 PF (PENN/) PENN S G.
 XX (RANK/) RANK D R.
 PA (HANK/) HANZEL D K.
 XX
 PI -Penn SG, Rank DR, Hanzel DK;
 XX WPI; 2004-119264/12.
 DR
 XX New human genome-derived single exon nucleic acid probes useful for human
 PT gene expression analysis, for identifying or characterizing alternative
 PT splicing events, for assessing genomic alterations or as tools for
 PT surveying tissues.
 XX
 PS Claim 15; SEQ ID NO 10603; 80pp; English.
 XX
 CC The invention relates to a nucleic acid probe for measuring human gene
 CC expression, comprising any of the 27,400 fully defined nucleotide
 CC sequences in the specification, or their complements or fragments, and
 CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
 CC fully defined in the specification. The probe is a single exon probe that
 CC hybridises under high stringency conditions to a nucleic acid molecule
 CC expressed in human cells or tissues. Also included are a spatially-
 CC addressable set of single exon nucleic acid probes for measuring human
 CC gene expression (comprising a plurality of single exon nucleic acid
 CC probes cited above, where each of the plurality of probes is separately
 CC and addressable or amplifiable from the plurality), a single
 CC exon microarray for measuring human gene expression, a method of
 CC measuring human gene expression, a vector comprising the single exon
 CC probe cited above, an ORF-encoded peptide comprising at least 8
 CC contiguous amino acids of any of the above-mentioned amino acid
 CC sequences (optionally with conservative amino acid substitutions), an
 CC isolated antibody that binds specifically to a peptide cited above,
 CC a method of selling and/or licensing single exon probes or microarrays to
 CC a customer desiring to measure gene expression, a method of providing
 CC human gene expression data by subscription, and a computer-readable
 CC storage medium which contains a database having a plurality of records
 CC (each record including data on the expression of a single exon probe
 CC cited above). The probe, methods and apparatus are useful in gene
 CC expression analysis. The probes may be used as tools for surveying
 CC tissues to detect the presence of expressed messages that contain their
 CC specific exon, or in constructing genome-derived single exon microarrays.
 CC In addition, the probes are used in identifying and characterising
 CC alternative splicing events, in detecting and characterising gross
 CC alterations in the genomic locus that includes their exon, in assessing
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,
 CC or in expressing the ORF-encoded peptide. The present sequence is a human
 CC single exon probe of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030194704
 XX
 SQ Sequence 531 BP; 117 A; 153 C; 149 G; 112 T; 0 U; 0 Other;

Query Match 20.9%; Score 233.2; DB 12; Length 531;
 Best Local Similarity 98.7%; Pred. No. 8e-57;
 Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 758 CAAGGTGGACCGAGCGGAACGTTTGGCCCCATTTGGCAGCGAGAGGAACTTACCT 817
 Db 34 CCAGGTGGACCGAGCGGAACGTTTGGCCCCATTTGGCAGCGAGAGGAACTTACCT 93
 QY 818 GATTGACTTGGCTGGGTCAGAGGACACCGCGGCGCAGGCAACAGGCGCTTCGGCTAAA 877
 Db 94 GATTGACTTGGCTGGGTCAGAGGACACCGCGGCGCAGGCAACAGGCGCTTCGGCTAAA 153

QY 878 AGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAGTGGTAGATCGCGCTGAA 937
DB 154 AGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAGTGGTAGATCGCGCTGAA 213
QY 938 TCAGGGGCTCCCTCGTGTACCTTATCGGGACACGACGCTACCTGCGCTATTGCGAGAC 995
DB 214 TCAGGGGCTCCCTCGTGTACCTTATCGGGACACGACGCTACCTGCGCTATTGCGAGTC 271

RESULT 14
ACH91108
ID ACH91108 standard; DNA; 232 BP.
XX
AC ACH91108;
XX
XX 29-JUL-2004 (first entry)
XX Human genome derived single exon probe #24303.
XX Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN//) PENN S G.
PA (RANK//) RANK D R.
PA (HANZ//) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
XX WPI; 2004-119264/12.

XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 1; SEQ ID NO 24303; 80pp; English.
XX
XX The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 688 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridizes under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.

CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
XX

XX Sequence 232 BP; 55 A; 64 C; 68 G; 45 T; 0 U; 0 Other;

Query Match 20.7%; Score 230.4; DB 12; Length-232;

Best Local Similarity 99.6%; Pred. No. 3.6e-56;

Matches 231; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 762 GTGACACGAGCGGGAACGTTTGGCCCCATTTCGCCAGCGAGAGGAAAACTTACCTGATT 821
DB 1 GTGACACGAGCGGGAACGTTTGGCCCCATTTCGCCAGCGAGAGGAAAACTTACCTGATT 60
QY 822 GACTTGGCTGGGTCTAGAGGACMACCGGGGCACAGGCAACAGGCGCTTCGGCTAAAGAG 881
DB 61 GACTTGGCTGGGTCTAGAGGACMACCGGGGCACAGGCAACAGGCGCTTCGGCTAAAGAG 120
QY 882 AGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAGTGGTAGATCGCGCTGAATCAG 941
DB 121 AGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAGTGGTAGATCGCGCTGAATCAG 180
QY 942 GGCTCCCTCGTGTACCTTATCGGACAGCAAGCTCAGCTGCGCTATTGCGAGG 993
DB 181 GGCTCCCTCGTGTACCTTATCGGACAGCAAGCTCAGCTGCGCTATTGCGAGG 232

RESULT 15

ADM01564

ID ADM01564 standard; cDNA; 2095 BP.

XX AC ADM01564;

XX 20-MAY-2004 (first entry)

XX Human cDNA of the invention SEQ ID NO:249.

XX ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.

XX Homo sapiens.

XX EP1347046-A1.

XX 24-SEP-2003.

XX 12-APR-2002; 2002EP-00008400.

XX 22-MAR-2002; 2002JP-00137785.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

XX Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX WPI; 2003-723558/69.

XX P-PSDB; ADM04007.

XX New polynucleotides and polypeptides are useful in gene therapy, for

XX developing a diagnostic marker or medicines for regulating their

XX expression and activity, or as a target of gene therapy.

XX Claim 1; SEQ ID NO 249; 305pp; English.

XX The invention relates to a novel human polynucleotide and the encoded

XX polypeptide. A polynucleotide of the invention may have a use in gene

XX therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful

CC as a primer for synthesizing the polynucleotide or as a probe for
 CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
 CC useful in gene therapy, for developing a diagnostic marker or medicines
 CC for regulating their expression and activity, or as a target of gene
 CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
 CC are useful as pharmaceutical agents. The present sequence represents a
 CC cDNA sequence of the invention.

XX
 SQ Sequence 2095 BP; 489 A; 603 C; 658 G; 345 T; 0 U; 0 Other;

Query Match 16.6%; Score 184.6; DB 11; Length 2095;
 Best Local Similarity 53.7%; Pred. No. 1.6e-42;
 Matches 463; Conservative 0; Mismatches 379; Indels 21; Gaps 3;

QY	264	TACCAAGAGCCTCATCGAGGCGCTCATCTGAGGAGAGGAGTACTCAGCAGGACATCTATGCAAGTTCA	323
DB	286	TACCTGTTGAGCGTGGCCCTTTGACTTCACCGCCACCAGGAGATGGTGTATCAGGCCACC	345
QY	324	GTGACGCCCATCTTAAGGCATCTTGCTGGAAGGCGCAGATGCCAGTGTGCTTCCCTATGGA	383
DB	346	ACCAAGAGCCTCATCGAGGCGCTCATCTCAGGCTCAATGCCACTGTCTTTGCCCTATGSC	405
QY	384	CCACAGAGAGCTGGGAGAGCAGCAGCATGCTGGGCGGCCAGAGCAACCTGGGGTGATC	443
DB	406	CCACAGAGCTGGGGAACCTACACCATGCTGGGCGCAGACCAGGAGCCTGGCATCTAT	465
QY	444	CGCGGGCTCTCATGACCTCTGCGAGCTCACAGGGAGGAGGCTGCCGAGGGCGGCCA	503
DB	466	GTTGAGACCTTCAGACCTCTTCGTGCCATCGAGGAGACCAAGCATGA-----C	516
QY	504	TGGGCGCTTTGTGCCCATGCTTACCTAGAGATCTACAGGAGAAGGTATTAGACCTC	563
DB	517	ATGGAGTATAGGTCTCCATGCTCTACCTGGAGATCTCAATGAGATGATCCGGGACCTG	576
QY	564	CTGGACCTCTGCTCGGAGACCTGTGTAATCCAGAGAGCTGCCGGGGGAATCTCTGATT	623
DB	577	TGAAACCCCTCCCTGGGCTACCTGGAGCTGCGGGAGGACTCTAAGGGGGGTGATCCAGGTG	636
QY	624	CGGGTCTCTCCAGAACCCCATAGTAGCTTTGTGATTTTGGCGGCACCTTCCTGCCA	683
DB	637	GCCGCATCACCGAAGTCTCCACCATCATGCAAGGAGATCATGAGCTGCTGATGAAG	696
QY	684	GCCAGTCGAATCGGACTGTAGAGCCACCCGGCTCAACAGCGCTCCTCCGCGAGTCA	743
DB	697	GGGAACCGCAGAGGACCCAGGAGCCACCGCGGCCAACCCAGAGCTCTCCGCTCCAC	756
QY	744	GCTGTCTCTGCTCAAGGTGACAGCGGGAAAGTTTGGCCCCATTTGCCA-----G	797
DB	757	GCGGTACTGCAAGGTGACCGTGCAGCGCAGCCGGGTCAAGAACATCTTGCAGGAGGTG	816
QY	798	CGAGAGGGAATACTCTACTGATTGACTTTGGCTGGGTGAGAGCAACCGGCGCACAGGC	857
DB	817	CGGAGGGCGCGCTGTTTATGATGACCTGGCTGGCTCAGAGCGGCTTCGAGACACAG	876
QY	858	AACAAGGCGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCTGGGC	917
DB	877	AATCGTGGCAGCGTATGAAGAGGGGGCCACATCAACCGCTCACTGTGSCACTGGGC	936
QY	918	AAAGTGTAGATCGCGCTG-----AATCAGGGCCTCCCTCGTGTACCTTATCGGGACAGC	971
DB	937	AACTGCATCAAGCGCTTGAGCGACAAAGGGTAGCAACAAGTACATCAACTATCGCGACAGC	996
QY	972	AAGCTCACTCGGCTATTGAGGACTCTCTGGGTGGCTCAGCCACAGTATCCTTATTGCC	1031
DB	997	AAGCTCACCGGCTCCTGAAGGACTCTCTGGGAGGAAACAGCCGACAGTATGATCGCT	1056
QY	1032	AACATTGCCCTGAGAGAGCGCTTCTTACCTTAGACACAGTCTCCGCACTCAACTTTTGTGCC	1091
DB	1057	CACATCAGTCTCGAGCAGTGCCTTCGAGGAGTCCCGGAACACCTGACCTACGCCGGC	1116
QY	1092	AGGTCAAAGGAGGTGATCAATTG	1114
DB	1117	CGGGCCAAAGACATTAAAGACTAG	1139

Search completed: November 10, 2004, 06:57:50
 Job time : 582.523 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 04:25:02 ; Search time 5089.37 Seconds
(without alignments)
10482.189 Million cell updates/sec

Title: US-10-797-893-7
Perfect score: 1464
Sequence: 1 atggctgctgtcggtgaag.....accattgtcccaaatgta 1464

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_hc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gse1.*
9: gb_gse2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1454.8	99.4	2016	3	CR590954
2	1454.8	99.4	2025	3	CR590974
3	1454.8	99.4	2057	3	CR615781
4	1067.8	72.9	2055	3	AK084310
5	1063.6	72.7	2056	3	AK075922
6	1007.6	68.8	1124	1	AL545712
7	973.8	66.5	1027	5	BX337406
8	970.6	66.3	1032	5	BX337353
9	922.4	63.0	1031	4	BM910587
10	912.8	62.3	1083	1	AL559772
11	897.6	61.3	1016	1	AL518361
12	882.8	60.3	1034	5	BQ051730
13	881.6	60.2	959	5	BX363358
14	875.4	59.8	971	5	BQ053671
15	874.5	59.7	958	5	BX354166
16	866.2	59.2	978	5	EX400806
17	860.8	58.8	1094	5	EX376199
18	854.4	58.4	926	1	AL578338
19	852.5	58.2	1019	4	BM546877
20	851.4	58.2	865	6	CD109092
21	849.4	58.0	999	1	AL559769
22	841.6	57.5	915	5	EX383181
23	823.6	56.3	1139	4	BM553107
24	821.4	56.1	1640	3	AK088033

25	820.8	56.1	978	4	BI754589
26	819.4	56.0	981	4	BM561588
27	817.8	55.9	1051	4	BM806996
28	817.6	55.8	980	5	EX404364
29	817.2	55.8	1041	4	BM478215
30	813	55.5	908	5	BQ953445
31	808.4	55.2	878	1	AL576578
32	804.8	55.0	1040	4	BM546897
33	802.6	54.8	845	4	BG686012
34	795.6	54.3	865	1	AL555550
35	790.6	54.0	855	5	EX372162
36	789.6	53.9	991	5	EX333336
37	786	53.7	1103	4	BM458841
38	782.4	53.4	995	1	AL559752
39	782.2	53.4	905	5	EX387403
40	782	53.4	921	5	EU860011
41	773.8	52.9	913	5	EX445789
42	772	52.7	990	5	EX399192
43	768.6	52.5	846	5	EX404325
44	767.4	52.4	1142	4	BM918115
45	761	52.0	838	5	EX341048

ALIGNMENTS

RESULT 1
CR590954
LOCUS
DEFINITION
full-length cDNA clone CS0DI040YI06 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION
CR590954
VERSION
CR590954.1 GI:50471761
KEYWORDS
HTC; CNSLT_CDNA.
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2016)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
JOURNAL
Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 2016)
Genoscope.
AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES
source
1..2016
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI040YI06"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 99.4%; Score 1454.8; DB 3; Length 2016;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 GGTGCTGTGGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCAGCTCGCGTAAGG 63
|||||
Db 9 GGTGCTGTGGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCAGCTCGCGTAAGG 68
|||||

Qy	1144	GAGGAAGAGAGATGGGAGCCCTGAGCCCATGGCAGCTTCAGCCTCTGCCTCCAGAAA	1201
Db	1149	GAGGAAGAGAGATGGGAGCCCTGAGCCCATGGCAGCTTCAGCCTCTGCCTCCAGAAA	1208
Qy	1204	CTCAGCCCCCTACAGAGACTAAGCAGCATGAGCCCGGCCCATGCTGGAGCGCCTCCTCAGC	1263
Db	1209	CTCAGCCCCCTACAGAGACTAAGCAGCATGAGCCCGGCCCATGCTGGAGCGCCTCCTCAGC	1268
Qy	1264	TTGACCGCTCTGCTTGGCTCCCGAGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG	1323
Db	1269	TTGACCGCTCTGCTTGGCTCCCGAGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG	1328
Qy	1324	CGAGAGCGGATGCTGCTTAATGAAGACAGTGAAGAGAGAGACCTAGAGATTGAGAGCTT	1383
Db	1329	CGAGAGCGGATGCTGCTTAATGAAGACAGTGAAGAGAGAGACCTAGAGATTGAGAGCTT	1388
Qy	1384	AAGACGAAGCAAAAGAACTTGGAGGCCAAGATGTTGGCCCAAGAGGCTGAGGAAAAGAG	1443
Db	1389	AAGACGAAGCAAAAGAACTTGGAGGCCAAGATGTTGGCCCAAGAGGCTGAGGAAAAGAG	1448
Qy	1444	AACCATGTTCCCAATG 1461	
Db	1449	AACCATGTTCCCAATG 1466	
RESULT 2			
CR590974		2025 bp mRNA linear HTC 21-JUL-2004	
LOCUS		full-length cDNA clone CSODI009YJ01 of Placenta Cot 25-normalized	
DEFINITION		of Homo sapiens (human).	
ACCESSION		CR590974	
VERSION		CR590974.1 GI:50471781	
KEYWORDS		HTC; CNSLT_cDNA	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		Li,W.B., Gruber,C., Jesse,J. and Polayes,D.	
TITLE		Full-length cDNA libraries and normalization	
JOURNAL		Unpublished	
REMARK		Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue	
REFERENCE		2 (bases 1 to 2025)	
AUTHORS		Genoscope.	
TITLE		Direct Submission	
JOURNAL		Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr	
COMMENT		- Web : www.genoscope.cns.fr)	
		1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primer end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.	
FEATURES		Location/Qualifiers	
source		1..2025	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="CSODI009YJ01"	
		/tissue type="Placenta Cot 25-normalized"	
		/plasmid="pCMVSPORT_6"	
ORIGIN			
Query Match		99.4%; Score 1454.8; DB 3; Length 2025;	
Best Local Similarity		99.9%; Pred. No. 0;	
Matches 1456; Conservative		0; Mismatches 2; Indels 0; Gaps 0;	
Qy	4	GCTGCTGTCGCGCTAAGCAAGATTGGAGTACTCGTGCCTCCACCTCCAGCTCGCGTAAGG	63
Db	59	GCTGCTGTCGCGCTAAGCAAGATTGGAGTACTCGTGCCTCCACCTCCAGCTCGCGTAAGG	118

QY 64 GTGGCTGTGCGACTGCGGCCAATTGTGATGGAACAGCGGAGCAAGTATCCCCCTGT 123
DB 119 GTGGCTGTGCGACTGCGGCCAATTGTGATGGAACAGCGGAGCAAGTATCCCCCTGT 178
QY 124 GTGGGGGCATGGACAGCTGCTCTAGAGATTGCTAACTGAGGAACCAACAGAGACT 183
DB 179 GTGGGGGCATGGACAGCTGCTCTAGAGATTGCTAACTGAGGAACCAACAGAGACT 238
QY 184 CTCAAATACCAATTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGAGCATCTATGCA 243
DB 239 CTCAAATACCAATTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGAGCATCTATGCA 298
QY 244 GGTTCAGTGACGCCCATCTTAAGGCACTGTCTGGAAGGCGCAAGTCCAGTGTGCTTGCC 303
DB 299 GGTTCAGTGACGCCCATCTTAAGGCACTGTCTGGAAGGCGCAAGTCCAGTGTGCTTGCC 358
QY 304 TATGACCCACAGGAGCTGGGAAGCAGCACCAATGCTGGCAGCCGCCAGCAACCTGGG 363
DB 359 TATGACCCACAGGAGCTGGGAAGCAGCACCAATGCTGGCAGCCGCCAGCAACCTGGG 418
QY 364 GTGATCCCGCGGGCTCTCATGGAATCTCTGAGCTCAAGAGGAGGAGGTCCGAGGGC 423
DB 419 GTGATCCCGCGGGCTCTCATGGAATCTCTGAGCTCAAGAGGAGGAGGTCCGAGGGC 478
QY 424 CGGCATGGGCGCTTTCTGTCAACATGTCTTACCTAGAGATCTTACAGGAGAAGTATTA 483
DB 479 CGGCATGGGCGCTTTCTGTCAACATGTCTTACCTAGAGATCTTACAGGAGAAGTATTA 538
QY 484 GACCTCTGGACCTGCTTGGGAGACTGCTTAATCCGAGAGACTGCGGGGGAATATC 543
DB 539 GACCTCTGGACCTGCTTGGGAGACTGCTTAATCCGAGAGACTGCGGGGGAATATC 598
QY 544 CTGATTCGGGTCTCTCCAGAAGCCATCAGTAGCTTTGCTGATTTTGGCGGCACCTTC 603
DB 599 CTGATTCGGGTCTCTCCAGAAGCCATCAGTAGCTTTGCTGATTTTGGCGGCACCTTC 658
QY 604 CTGCGACAGTCGAATTCGAGTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGC 663
DB 659 CTGCGACAGTCGAATTCGAGTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGC 718
QY 664 AGTCATCTGTCTCTGCTCAAGGTGAGCCAGCGGGAACGTTTGGGCCCATTTTCGCAG 723
DB 719 AGTCATCTGTCTCTGCTCAAGGTGAGCCAGCGGGAACGTTTGGGCCCATTTTCGCAG 778
QY 724 CGAGAGGGAACCTCTACCTGATTGACTTGGCTGGTCCAGAGGACAAACGCGCACAGGC 783
DB 779 CGAGAGGGAACCTCTACCTGATTGACTTGGCTGGTCCAGAGGACAAACGCGCACAGGC 838
QY 784 AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTTGTCTGGGC 843
DB 839 AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTTGTCTGGGC 898
QY 844 AAGTGTAGATGGCTGAATCAGGGCTCTCTGCTGTAACCTTATCGGAGCAGCAAGCTC 903
DB 899 AAGTGTAGATGGCTGAATCAGGGCTCTCTGCTGTAACCTTATCGGAGCAGCAAGCTC 958
QY 904 ACTCGCTATTGACGAGTCTCTCGGTGGTCCAGCCACAGTATCTTTATTTGCCAAT 963
DB 959 ACTCGCTATTGACGAGTCTCTCGGTGGTCCAGCCACAGTATCTTTATTTGCCAAT 1018
QY 964 GCCCTGAGAGAGCTTTCTACCTAGACACAGTCTCCGACTCAACTTTGCTGCCAGTCC 1023
DB 1019 GCCCTGAGAGAGCTTTCTACCTAGACACAGTCTCCGACTCAACTTTGCTGCCAGTCC 1078
QY 1024 AAGAGTGTATCAATCGGCTTTTACCAATGAGAGCTTGAGCCTCATGCTTGGGACCT 1083
DB 1079 AAGAGTGTATCAATCGGCTTTTACCAATGAGAGCTTGAGCCTCATGCTTGGGACCT 1138
QY 1084 GTTAAGCTGTCTCAGAAGAATTTGCTTGGTCCACCAAGAGCAAGAGAGCCCGAGGCT 1143
DB 1139 GTTAAGCTGTCTCAGAAGAATTTGCTTGGTCCACCAAGAGCAAGAGAGCCCGAGGCT 1198
QY 1144 GAGGAAGAGGAGATTGGGAGCCCTGAGCCCATGCGAGCTCCAGCTCTGCTCTCCAGAAA 1203

DB 1199 GAGGAAGAGGAGATCGGGAGCCCTGAGCCCTGAGCCCTCTGCTCCAGAAA 1258
QY 1204 CTCAGCCCTTACAGAGCTAAGCAGCATGACCCGGCCATGCTGGAGGCTCTCTCAGC 1263
DB 1259 CTCAGCCCTTACAGAGCTAAGCAGCATGACCCGGCCATGCTGGAGGCTCTCTCAGC 1318
QY 1264 TTGACCGTCTGCTTGGCTCCAGGGAGCCAGGGGGCCCTCTCTGCTGAGTACCCCAAG 1323
DB 1319 TTGACCGTCTGCTTGGCTCCAGGGAGCCAGGGGGCCCTCTCTGCTGAGTACCCCAAG 1378
QY 1324 CGAGAGCGGATGCTGCTTAATGAAGCAGTGAAGAGAGGACCTTAGAGATTGAGAGGCTT 1383
DB 1379 CGAGAGCGGATGCTGCTTAATGAAGCAGTGAAGAGAGGACCTTAGAGATTGAGAGGCTT 1438
QY 1384 AAGACGAGCAAAAAGAACTGGAGCCCAAGTGTGGCCAGAGGCTGAGAGAAAGAG 1443
DB 1439 AAGACGAGCAAAAAGAACTGGAGCCCAAGTGTGGCCAGAGGCTGAGAGAAAGAG 1498
QY 1444 AACATTGTCCCAATG 1461
DB 1499 AACATTGTCCCAATG 1516

RESULT 3
CR615781
LOCUS 2057 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CSODK012YE23 of HeLa cells Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR615781
VERSION CR615781.1 GI:50496588
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2057)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
REFERENCE 2 (bases 1 to 2057)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source 1..2057
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODK012YE23"
/tissue types="HeLa cells Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 99.4%; Score 1454.8; DB 3; Length 2057;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 GTTCGCTGTGCGCTAAGCAGATTGGAGCTACTGCTGCTCCACTCCAGCTCGGTAAGG 63
DB 79 GTTCGCTGTGCGCTAAGCAGATTGGAGCTACTGCTGCTCCACTCCAGCTCGGTAAGG 138
QY 64 GTTCGCTGTGCGCTAAGCAGATTGGAGCTACTGCTGCTCCACTCCAGCTCGGTAAGG 123

Db	139	GTGGCTGTGCGAATCGGCGCAATTTGTGGATGGAACAGCGGGAGCAAGTGAATCCCGCCCTGT	198	
Qy	124	GTGCGGGGATGAGACAGCTGCTCTTAGAGATTGCTTAAGTGGAGAAACACAGGAGACT	183	
Db	199	GTGCGGGGATGAGACAGCTGCTCTTAGAGATTGCTTAAGTGGAGAAACACAGGAGACT	258	
Qy	184	CTCAATATACAGTTTGTATGATCCCTTCTATGGGAGAGAGGATCTCAGCAGGACATCTATGCA	243	
Db	259	CTCAATATACAGTTTGTATGATCCCTTCTATGGGAGAGAGGATCTCAGCAGGACATCTATGCA	318	
Qy	244	GGTTTCACTAGTCAGCCCATCTAAGGCACTTTGCTGGAAGGCGCAATTCGACGATGCTTGGC	303	
Db	319	GGTTTCACTAGTCAGCCCATCTAAGGCACTTTGCTGGAAGGCGCAATTCGACGATGCTTGGC	378	
Qy	304	TATGGACCCACAGGAGCTGGAGAGCGCACAAATGCTGGGAGCCCGCAGAGCAACTGGG	363	
Db	379	TATGGACCCACAGGAGCTGGAGAGCGCACAAATGCTGGGAGCCCGCAGAGCAACTGGG	438	
Qy	364	GTGATCCCGGGGCTCTCATGGACCTCTCTGAGCTCTCTGAGCTCAAGGAGGAGGCTCCGAGGGC	423	
Db	439	GTGATCCCGGGGCTCTCATGGACCTCTCTGAGCTCTCTGAGCTCAAGGAGGAGGCTCCGAGGGC	498	
Qy	424	CGGCCATGGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTACAGGAGAGGTATTA	483	
Db	499	CGGCCATGGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTACAGGAGAGGTATTA	558	
Qy	484	GACCTCTGAGACCTGCTCGGAGAGCTGTTAATCCGAGAAGACTGCGGGGGGATATC	543	
Db	559	GACCTCTGAGACCTGCTCGGAGAGCTGTTAATCCGAGAAGACTGCGGGGGGATATC	618	
Qy	544	CTGATTCGGGGTCTCTCCAGAGACCCATCAGTAGCTTTGCTGATTTAGCGGCACTTC	603	
Db	619	CTGATTCGGGGTCTCTCCAGAGACCCATCAGTAGCTTTGCTGATTTAGCGGCACTTC	678	
Qy	604	CTGCCACCCAGTCGAATTCGATGTAGGAGCCACCGGCTCAACAGAGGCTCTCCCGC	663	
Db	679	CTGCCACCCAGTCGAATTCGATGTAGGAGCCACCGGCTCAACAGAGGCTCTCCCGC	738	
Qy	664	AGTCATGCTGTGCTCTCTGTCAGAGTGGAGCCAGCGGAACTTTGGCCCCCAATTGCGCAG	723	
Db	739	AGTCATGCTGTGCTCTCTGTCAGAGTGGAGCCAGCGGAACTTTGGCCCCCAATTGCGCAG	798	
Qy	724	CGAGAGGGAATCTACCTGATTTGATCTTGGCTGGGTGAGAGCAACCGGCGCACAGGC	783	
Db	799	CGAGAGGGAATCTACCTGATTTGATCTTGGCTGGGTGAGAGCAACCGGCGCACAGGC	858	
Qy	784	AACAAGGCGCTTCGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGC	843	
Db	859	AACAAGGCGCTTCGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGC	918	
Qy	844	AAAGTGTGATGCTGCTGAATCAGGCGCTCCCTGCTGTAACCTATCGGACAGCAAGCTC	903	
Db	919	AAAGTGTGATGCTGCTGAATCAGGCGCTCCCTGCTGTAACCTATCGGACAGCAAGCTC	978	
Qy	904	ACTGCTTATTCAGGAGCTCTGCTGGTGGCTCAGCCACATGATCTTTATTGCCAACATT	963	
Db	979	ACTGCTTATTCAGGAGCTCTGCTGGTGGCTCAGCCACATGATCTTTATTGCCAACATT	1038	
Qy	964	GCCCTCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTCCAGGTCC	1023	
Db	1039	GCCCTCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTCCAGGTCC	1098	
Qy	1024	AAGAGGTGATCAATTCGGCTTTTACCAATGAGAGCTGAGCCCTCATGCTTGGGACCT	1083	
Db	1099	AAGAGGTGATCAATTCGGCTTTTACCAATGAGAGCTGAGCCCTCATGCTTGGGACCT	1158	
Qy	1084	GTTAGAGTCTCTCAGAGAGATTGCTTGGTCCACAGAGGCAAGAGAGCCCGAGGCCCT	1143	
Db	1159	GTTAGAGTCTCTCAGAGAGATTGCTTGGTCCACAGAGGCAAGAGAGCCCGAGGCCCT	1218	
Qy	1144	GAGGAGAGGAGATTGGAGGCCCTGAGCCCATGAGGAGCTCGAGCTCTGCTCCCGAGAA	1203	
Db	1219	GAGGAGAGGAGATTGGAGGCCCTGAGCCCATGAGGAGCTCGAGCTCTGCTCCCGAGAA	1278	
Qy	1204	CTCAGCCCCCTACAGAGCTTAAGCAGCATGGAACCGGCCCATGCTGGAGCGCCTCTCAGC	1263	
Db	1279	CTCAGCCCCCTACAGAGCTTAAGCAGCATGGAACCGGCCCATGCTGGAGCGCCTCTCAGC	1338	
Qy	1264	TTGAGACGCTGCTGCTTCCCTCCAGGGGAGCGAGGGGCCCTCTGTTTGTAGTACCCCAAG	1323	
Db	1339	TTGAGACGCTGCTGCTTCCCTCCAGGGGAGCGAGGGGCCCTCTGTTTGTAGTACCCCAAG	1398	
Qy	1324	CGAGAGCGGATGCTGCTTAATGAAGACAGTAGTAAGAGAGAGGACCTTAGAGATTGAGGCTT	1383	
Db	1399	CGAGAGCGGATGCTGCTTAATGAAGACAGTAGTAAGAGAGAGGACCTTAGAGATTGAGGCTT	1458	
Qy	1384	AAGACGAGCAAAAGAACTGGAGGCCCAAGATGTTGCCCGCAGAGGCTGAGGAAAAAGAG	1443	
Db	1459	AAGACGAGCAAAAGAACTGGAGGCCCAAGATGTTGCCCGCAGAGGCTGAGGAAAAAGAG	1518	
Qy	1444	AACCATTTGCTCCCAATG 1461		
Db	1519	AACCATTTGCTCCCAATG 1536		
RESULT 4				
AK084310				
LOCUS				
DEFINITION				
AK084310 2055 bp mRNA linear HTC 03-APR-2004				
Mus musculus 12 days embryo eyeball cDNA, RIKEN full-length				
enriched library, clone:D230020105 product:kinesin superfamily				
protein KIF22, full insert sequence.				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
1 Carninci, P. and Hayashizaki, Y.				
High-efficiency full-length cDNA cloning				
Meth. Enzymol. 303, 19-44 (1999)				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,				
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
Normalization and subtraction of cap-trapper-selected cDNAs to				
prepare full-length cDNA libraries for rapid discovery of new genes				
Genome Res. 10 (10), 1617-1630 (2000)				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,				
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,				
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,				
Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,				
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,				
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Kira, A. and Hayashizaki, Y.				
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.				
RIKEN integrated sequence analysis (RISA) system--384-format				
sequencing pipeline with 384 multicapillary sequencer				
Genome Res. 10 (11), 1757-1771 (2000)				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
4 The RIKEN Genome Exploration Research Group Phase II Team and the				
FANTOM Consortium.				
Functional annotation of a full-length mouse cDNA collection				
Nature 409, 685-690 (2001)				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
5 The FANTOM Consortium and the RIKEN Genome Exploration Research				
Group Phase I & II Team.				
Analysis of the mouse transcriptome based on functional annotation				
of 60,770 full-length cDNAs				

Fri Nov 12 12:26:26 2004

us-10-797-893-7.rst

cdna library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Sciences Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>
URL: <http://fantom.gsc.riken.jp/>
Location/Qualifiers
1. .2056
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:2400004116"
/db_xref="taxon:10090"
/clone="2400004116"
/cell_type="ES cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
misc_feature
1. .2056
/note="kinesin superfamily protein KIF22"

FEATURES
source

COMMENT

ORIGIN

Query Match 72.7%; Score 1063.6; DB 3; Length 2056;
Best Local Similarity 83.9%; Pred. No. 1.5e-272; Indels 0; Gaps 0;
Matches 1201; Conservative 0; Mismatches 230;

QY 6 TCGCTGTCGGCTAAGCAAGATTGGAGCTACTCTGTCGTCACCTCCAGCTCCGCTAAGGGT 65
DB 73 TGGCGCTGTGTAGCAAGAGGGCTTGGCGGGTCCGCTCTAGCTGAGTAAGGGT 132

QY 66 GGCTGTGCGACTCGCGCCATTGTGGATGGAAAGCGGGCAAGTATCCCCCTGTGT 125
DB 133 GGCTGTGCGCTGGCGCCGCTTATGGATGGAGACAGAGGAGCTCCCGCGTGT 192

QY 126 GCGGGGATGACAGCTCTCTAGGATTTCTAACTGGAGGAACACACAGGAGACTCT 185
DB 193 CGAGCCATAGACAGCTCTCTTAAAGTGGCTAACTGGAGGAATACACAGGAGCGT 252

QY 186 CAAATACCAAGTTTGATGCTTCTATGGGGAGAGGATGACTCAGCAGGACATCTATGAGG 245
DB 253 CAAATATCAGTTTGATGCTTCTATGGCGAGAGGACACTCAGCAGGAGTCTATGTAGG 312

QY 246 TTCACTGACGCCCATCTTAAGCACTTCTGGAGGGCAGAGTCCAGTGTGTCCTA 305
DB 313 TTCACTGACGCCCATCTTAAGCACTTCTGGAGGGCAGAGTCCAGTGTGTCCTA 372

QY 306 TGGACCCACAGGAGCTGGGAAGACGACACATGCTGGGGCAGCCAGAGCACTTGGGT 365
DB 373 TGGGGCTTCTGGGGCAGGAGGACACACATGCTGGGCAGCCAGAGCACTTGGGT 432

QY 366 GATCCCGGGCTCTCATGGACCTCTGAGCTCAAGGGAGGAGGTCCGAGGGCG 425
DB 433 GATTCCTGGGCTCTCATGGACCTCTGAGCTCAAGGGAGGAGGTCCGAGGGCG 492

QY 426 GCATCGGGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTATTAGA 485
DB 493 GCATCGGGAGCTTCTGTCATGCTTATTTAGAGATCTACAGGAGAGGTATTAGA 552

QY 486 CCTCTCGACCCCTGCTTCGGGAGACCTGCTTAATCCGAGAGACTGCGGGGGAATATCT 545
DB 553 CCTCTTGGATCTGTCATCAGGAGACCTGCTGATCCGGAAGACTGCGGGGGAATATCT 612

QY 546 GATTCGGGTCTCTCCAGAGCCCATCAGTCTTGTGATTTTGGCGGCACCTTCT 605
DB 613 GATTCGGGTCTCTCCAGAGCCCATCAGTCTTGTGATTTTGGCGGCACCTTCT 672

QY 606 GCCAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGTCTCTCCGAG 665
DB 673 TCACAGCCAGTAGAATTCGACTGTAGGAGCCACCCGGCTTAAACAGCGTCTCTCCGAG 732

QY 666 TCATGCTGCTCTCTGCTCAAGTGGACACGCGGAGACGTTTGGCCCCATTTCCGAGCG 725
DB 733 TCACGCGAGTCTCTTGGTCAAGGTAGATCAGCGGTGAACGTTTGACTCCATTTCCGAGCG 792

RESULT 5
AKO75922
LOCUS
DEFINITION
AKO75922 2056 bp mRNA linear HTC 03-APR-2004
Mus musculus ES cells cDNA, RIKEN full-length enriched library,
clone:2400004116 product:kinesin superfamily protein KIF22, full
insert sequence.
AKO75922 1 GI:26096579
HTC; CAP trapper.
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED
10349636
REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsui, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
PUBMED
11076861
REFERENCE
3 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2056)
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, J., Komori, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saio, H., Saio, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tegami, M., Tagawa, A., Takahashi, P., Tanaka, T.,
Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A.,
Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

QY 726 AGAGGAACTCTACCTGATTTGACTTGGCTGGGTTCAGAGGACCAACCGCGCACAGGCAA 785
 DB |||||
 QY 793 GGAAGGAACTCTACATTAATTGATTTGGCTGGTTCAGAGGACCAACCGTGCACAGGCAA 852
 DB |||||
 QY 786 CAAGGGCTTCGGCTTAAAGAGAGTGGAGCCATCAACACTCCCTTTGTTGCTTGGGCAA 845
 DB |||||
 QY 853 CAAGGGCTTCGGCTTAAAGAGAGCGGAGCCATCAACACTCCCTTTGTTGCTTGGGCAA 912
 DB |||||
 QY 846 AGTGTAGATCGCTGAATCAGGGCTCCCTCTGTTACTTATCGGGACACCAAGCTCAC 905
 DB |||||
 QY 913 AGTGTAGATCGATTAACACAGGGCTCCCTTCGATACCATACCGGACACCAAGCTCAC 972
 DB |||||
 QY 906 TCGCTATTTGAGGACTCTCTGGGTGGCTCAGCCACAGTATCTTATTTGCAACATTCG 965
 DB |||||
 QY 973 TCGTCTGCTGAGGACTCTCTGGGAGGCTCAGCTCATAGCATCTCTATTGCAACATTCG 1032
 DB |||||
 QY 966 CCCTGAGAGAGCTTCTACCTAGACACAGTCTCGCATCACTCACTTTGCTGCGAGTCAA 1025
 DB |||||
 QY 1033 TCCTGGAAGAGCTTTTACCAAGGATTAACAATTCAGCATTTCACTGCTAGGTCAA 1092
 DB |||||
 QY 1026 GGAGTGATCAATCGGCTTTTACCAATGAGAGCTCGAGCTCATGCTTTGGGACCTGT 1085
 DB |||||
 QY 1093 GGAGTGATTAACCGGCTTTTACCAATGAGAGTCTCAGCTCATGCTTTGGACCTGT 1152
 DB |||||
 QY 1086 TAAGCTGTCTCAGAAAGAAATTTGTTGGTCCACAGAGGCAAGAGGCGCGGCTCTGA 1145
 DB |||||
 QY 1153 TAAGCTGTCTCAGAAAGAAATTTGTTGGTCCACAGAGGCAAGAGGCGCGGCTCTGA 1212
 DB |||||
 QY 1146 GGAAGAGAGATTTGGAGCCCTGAGCCATGCGAGCTTCAGCTCTGCTCCAGAACT 1205
 DB |||||
 QY 1213 AGAAGATCAACTGGGAGCCCTGAGTCTACAGAGCTCCAGCTCTGCTCCAGAACT 1272
 DB |||||
 QY 1206 CAGCCCTCTACAGAGCTTAAGCAGCATGAGCCCGGCTGCTGAGCGCTCTCAGCTT 1265
 DB |||||
 QY 1273 AAGCTCTTACAGAGCTTAAGCAGCATGAGTCTCAGCCATGTTGAGAACCTCTCAGTAT 1332
 DB |||||
 QY 1266 GGACCGTCTGCTTCCCTCCAGGGAGCCAGGGGCGCCCTCTGTTGAGTACCCCAAGCG 1325
 DB |||||
 QY 1333 GGAACGTTTGTGGGTTCCAGGAGCAAGGAGGACCCCTCTACTGAATACCCCAAGCG 1392
 DB |||||
 QY 1326 AGACCGATGTTGCTTAATGAAGACAGTAGAAGAGAGACCTAGAGATTGAGGCTTAA 1385
 DB |||||
 QY 1393 AGAACGAATGTTGCTCATGAGACAGTGGAGGAAAGAACTTAGAGATTGAGAGGCTTAA 1452
 DB |||||
 QY 1386 GACGAACAAAGAACTGGAGGCAAGATGTTGCGCCAGAGGCTGAGGA 1436
 DB |||||
 QY 1453 GATGAGCAAAAGAACTGGAAGCAAGGTTTGGCCAGAGGACCCGGA 1503
 DB |||||

RESULT 6
 AL545712
 LOCUS 1124 bp mRNA linear EST 25-MAR-2004
 DEFINITION AL545712 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 Clone CS0D1009YJ01 5-PRIME, mRNA sequence.
 ACCESSION AL545712
 VERSION AL545712.3 GI:45746191
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1124)
 Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:31267547.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 7766.r

For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?c=CS0D1009CE01QPI&c=7766.r.

FEATURES

source
 1. 1124
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1009YJ01"
 /tissue type="PLACENTA COT 25-NORMALIZED"
 /clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 68.8%; Score 1007; DB 1; Length 1124;
 Best Local Similarity 96.9%; Pred. No. 1.7e-257;
 Matches 101; Conservative 25; Mismatches 5; Indels 2; Gaps 2;
 QY 4 GGTGCTGTGGCTTAAGCAAGATTGGAGTACTGCTGCCACCTCCAGCTCGGCTAAGG 63
 DB |||||
 QY 59 GGTGCTGTGGCTTAAGCAAGATTGGAGTACTGCTGCCACCTCCAGCTCGGCTAAGG 118
 DB |||||
 QY 64 GTGCTGTGCACTGCGGCCATTGTTGGATGGAACAGCGGAGCAAGTATCCCCCTGT 123
 DB |||||
 QY 119 GTGCTGTGCACTGCGGCCATTGTTGGATGGAACAGCGGAGCAAGTATCCCCCTGT 178
 DB |||||
 QY 124 GTGCGGCGATGGACAGCTGCTCTAGAGATGCTTAAGTGGAGGACCAACAGGAGCT 183
 DB |||||
 QY 179 GTGCGGCGATGGACAGCTGCTCTAGAGATGCTTAAGTGGAGGACCAACAGGAGCT 238
 DB |||||
 QY 184 CTCAATATACAGTTTGTATGCTCTTATGGGAGAGGAGTACTACAGCAGGACATCTATGCA 243
 DB |||||
 QY 239 CTCAATATACAGTTTGTATGCTCTTATGGGAGAGGAGTACTACAGCAGGACATCTATGCA 298
 DB |||||
 QY 244 GGTTCAGTCAGCCCATCTTAAGGCACTTGTGGAAGGCGAGAAATGCCAGTGTCTTGC 303
 DB |||||
 QY 299 GGTTCAGTCAGCCCATCTTAAGGCACTTGTGGAAGGCGAGATTGTCACTGTCTTGC 358
 DB |||||
 QY 304 TATGCAACACAGAGCTGGGAGAGCCACACATGCTGGCGACGCCAGAGCACTCTGGG 363
 DB |||||
 QY 359 TATGCAACACAGAGCTGGGAGAGCCACACATGCTGGGAGCCACAGCAACCTGGG 418
 DB |||||
 QY 364 GTGATCCCGCGGCTCTCATGGACCTCTGAGCTCAAGAGGAGGAGGTCGCGAGGGC 423
 DB |||||
 QY 419 GKATCCCGCGGCTCTCATGGACCTCTGAGCTCAAGAGGAGGAGGTCGCGAGGGG 478
 DB |||||
 QY 424 CGGCGATGGCCCTTTCTGTCACTGCTTACCTAGAGATCTTACAGGAGAAAGTATTA 483
 DB |||||
 QY 479 CGGCGATGGCCCTTTCTGTCACTGCTTACCTAGAGATCTTACAGGAGAAAGKATTA 538
 DB |||||
 QY 484 GACCTCTGACCTCTGCTGGGAGAGCTGTAATCGAGAGAGCTCCCGGGGAAATATC 543
 DB |||||
 QY 539 GACCTCTGACCTCTGCTGGGAGAGCTGTAATCGAGAGAGCTCCCGGGGAGATATC 598
 DB |||||
 QY 544 CTGATTCGCGGTCTCTCCAGAGACCCATCAGTAGCTTTGCTGATTTTGGCGGCACCTTC 603
 DB |||||
 QY 599 CTGATTCGCGGTCTCTCCAGAGACCCATCAGTAGCTTTGCTGATTTTGGCGGCACCTTC 658
 DB |||||
 QY 604 CTGCGAGCCAGTGAATAATCGGATGTAGAGAGCCACCGGCTCAACAGCGCTCTCTCCGC 663
 DB |||||
 QY 659 CTGCGAGCCAGTGAATAATCGGATGTAGAGAGCCACCGGCTCAACAGCGCTCTCTCCGC 718
 DB |||||
 QY 664 AGTCATGCTGTCTCTGCTCAAGGTGGACAGGGGAAAGTGTGGGCCCATTTGGCGAG 723
 DB |||||
 QY 719 AGTCATGCTGTCTCTGCTCAAGGTGGACAGGGGAGCGTGTGGGCCCATTTGGCGAG 778
 DB |||||
 QY 724 CGAGAGGGAATACTCTACCTGATTGACTTGGCTGGGTTCAGAGGAGCAACCGCGCACAGGC 783
 DB |||||

Db	779	CGAGAGGGGAAACTTACCTGATTGATTTGGCTGGKCAAGGRCACCGGCGACAGGC	838
Qy	784	AAACAGGGCTTCGGCT-AAAAGAGAGTGAGGCATCAACACCTCCCTGTTTGTCTGGG	842
Db	839	AAACAGGGCTTCGGGTAAGAGAGAGTGAGGCATCAACACCTCCCTGTTTGTCTGGG	898
Qy	843	CAAAAGTGTAGATGCGCTGAATCAGGGCTCCCTGTTGATCTTATCGGACAGCAAGCT	902
Db	899	CAAAAGTGTAGATGCGCTGAATCAGGGCTCCCTGTTGATCTTATCGGACAGCAAGCT	958
Qy	903	CACCTCCCTATTGACAGACTCTCTGGGTGGCTCAGCCACAGTATCTTATGCGCAACAT	962
Db	959	CACCTCCCTATTGACAGACTCTCTGGGTGGCTCAGCCACAGTATCTTATGCGCAACAT	1017
Qy	963	TGCCCCCTGAGAGCGCTTCTACCTAGACAGTCTCCGACACTCACTTGTGCGCAGTTC	1022
Db	1018	TGCCCCCTGAGAGCGCTTCTACCTAGACAGTCTCCGACACTCACTTGTGCGCAGTTC	1077
Qy	1023	CAAGGAGGTGATCAATCGGCTTTTACCA	1051
Db	1078	CAAGGGGGKGRTCATCGGCTTTTACAA	1106
RESULT 7			
BX337406			
LOCUS			
DEFINITION			
BX337406 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA			
Clone CSOD1041Y113 5-PRIME, mRNA sequence.			
ACCESSION			
BX337406			
KEYWORDS			
EST.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1 (bases 1 to 1027)			
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.			
Full-length cDNA libraries and normalization			
Unpublished (2001)			
On May 2, 2003 this sequence version replaced gi:30341614.			
COMMENT			
Contact: Genoscope			
Genoscope - Centre National de Sequencage			
BP 191 91006 EVRY cedex - France			
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime			
end enriched, double-strand cDNA was digested with Not I and cloned			
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library			
was normalized. Library was constructed by Life Technologies, a			
division of Invitrogen. This sequence belongs to sequence cluster			
7766.r			
For more information about this cluster, see			
http://www.genoscope.cns.fr/cdna?c=CSOD1041AE07QF1&c=7766.r.			
FEATURES			
source			
1..1027			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="CSOD1041Y113"			
/tissue_type="PLACENTA COT 25-NORMALIZED"			
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"			
/note="1st strand cDNA was primed with a NotI-oligo(dT)			
primer. Five prime end enriched, double-strand cDNA was			
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."			
ORIGIN			
Query Match			
Best Local Similarity 99.6%; Pred. No. 1.2e-248;			
Matches 983; Conservative 3; Mismatches 0; Indels 1; Gaps 1;			
4 GGTGCGTGTGGCTAAGCAAGATTGGAGCTACTGTGTCGTCACCTCCAGCTCGCGTAGG			
Qy	4	GGTGGCTGTGGCTAAGCAAGATTGGAGCTACTGTGTCGTCACCTCCAGCTCGCGTAGG	63
RESULT 8			
BX337353			
LOCUS			
DEFINITION			
BX337353 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA			
clone CSOD1040Y106 5-PRIME, mRNA sequence.			
ACCESSION			
BX337353			
1032 bp mRNA linear EST 07-APR-2004			
Db	42	GGTGGCTGTGGCTAAGCAAGATTGGAGCTACTGTGTCGTCACCTCCAGCTCGCGTAGG	101
Qy	64	GTGGCTGTGGCACTGGGGCCATTGTTGGATGGAACAGCGGAGCAAGTGTATCCCTGT	123
Db	102	GTGGCTGTGGCACTGGGGCCATTGTTGGATGGAACAGCGGAGCAAGTGTATCCCTGT	161
Qy	124	GTGGGGGGCATGAGCAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACT	183
Db	162	GTGGGGGGCATGAGCAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACT	221
Qy	184	CTCAAAATACCAAGTTGATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA	243
Db	222	CTCAAAATACCAAGTTGATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA	281
Qy	244	GGTTCACTGAGAGCCATCTTAAGGCACTTCTGGAAGGGCAGATGCGAGTGTCTTGCC	303
Db	282	GGTTCACTGAGAGCCATCTTAAGGCACTTCTGGAAGGGCAGATGCGAGTGTCTTGCC	341
Qy	304	TATGAGCCCAAGAGAGTGGGAGACGACACAATGCTGGGAGAGGAGTGTCTTGCC	363
Db	342	TATGAGCCCAAGAGAGTGGGAGACGACACAATGCTGGGAGAGGAGTGTCTTGCC	401
Qy	364	GTGATCCCGCGGGCTCTCANGGACCTCTGAGCTCAAGGGAGAGGAGTGTCTTGCC	423
Db	402	GTGATCCCGCGGGCTCTCANGGACCTCTGAGCTCAAGGGAGAGGAGTGTCTTGCC	461
Qy	424	CGGCAATGGGCGCTTCTGTCAACCATGCTTACCTAGAGATCTACCAAGGAGAGTATTA	483
Db	462	CGGCAATGGGCGCTTCTGTCAACCATGCTTACCTAGAGATCTACCAAGGAGAGTATTA	520
Qy	484	GACCTCTCTGAGACCTCTGCTTGGGAGACCTGTAATCCGAGAAGACTGCGGGGGGATATC	543
Db	521	GACCTCTCTGAGACCTCTGCTTGGGAGACCTGTAATCCGAGAAGACTGCGGGGGGATATC	580
Qy	544	CTGATTCGGGCTCTCTCCAGAGCCATCAGTACTTGTGATTTTGGCGGCACTTC	603
Db	581	CTGATTCGGGCTCTCTCCAGAGCCATCAGTACTTGTGATTTTGGCGGCACTTC	640
Qy	604	CTGCGAGCCAGTCGAAATCGGACTCTAGGAGCCACCCGGCTCAACAGCGCTCTCCCGC	663
Db	641	CTGCGAGCCAGTCGAAATCGGACTCTAGGAGCCACCCGGCTCAACAGCGCTCTCCCGC	700
Qy	664	AGTCATGCTGCTCTCTGTCGAGTGAACAGGGGAACTGTTGGCCCCATTTGCGCAG	723
Db	701	AGTCATGCTGCTCTCTGTCGAGTGAACAGGGGAACTGTTGGCCCCATTTGCGCAG	760
Qy	724	CGAGAGGAAACTCTTACCTGATTCAGTCTGGCTCAGAGCAACCGGGGCGACAGGC	783
Db	761	CGAGAGGAAACTCTTACCTGATTCAGTCTGGCTCAGAGCAACCGGGGCGACAGGC	820
Qy	784	AAACAGGGCTTCGGGCTAAGAGAGTGGAGCCATCAACCTCTCTGTTGTTCTTGCGC	843
Db	821	AAACAGGGCTTCGGGCTAAGAGAGTGGAGCCATCAACCTCTCTGTTGTTCTTGCGC	880
Qy	844	AAAGTGGTAGATGCGCTGAATCAGGGCCCTCCCTCGGTACCTTATCGGACAGCAAGCTC	903
Db	881	AAAGTGGTAGATGCGCTGAATCAGGGCCCTCCCTCGGTACCTTATCGGACAGCAAGCTC	940
Qy	904	ACTCGCTATTGAGGACTCTCTGGGTGGCTCAGGCCACAGTATCTTATGCGCAACT	963
Db	941	ACTCGCTATTGAGGACTCTCTGGGTGGCTCAGGCCACAGTATCTTATGCGCAACT	1000
Qy	964	GCCCTCTGAGAGAGCTTCTACCTAGAC	990
Db	1001	GCCCTCTGAGAGAGCTTCTACCTAGAC	1027

/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN		Query Match	63.0%;	Score 922.4;	DB 4;	Length 1031;
		Best Local Similarity	99.0%;	Pred. No. 6.7e-235;		
		Matches	937;	Conservative	0;	Mismatches 8; Indels 1; Gaps 1;
Qy	103	GGAGCAAGTATCCCTCCCTGTGTGGGGCATGGACAGCTGCTCTAGAGATTGCTAAC	162			
Db	1	GGAGCAAGTATCCCTCCCTGTGTGGGGCATGGACAGCTGCTCTAGAGATTGCTAAC	60			
Qy	163	TGAGGAACCCACAGGAGACTCTCAATACCAATGATGCTTCTATGGGAGAGAGT	222			
Db	61	TGAGGAACCCACAGGAGACTCTCAATACCAATGATGCTTCTATGGGAGAGAGT	120			
Qy	223	ACTCAGCAGCATCTATGCAAGTTCAGTGCAGCCCATCTTAAGGCACCTTGTGGAAGG	282			
Db	121	ACTCAGCAGCAGCATCTATGCAAGTTCAGTGCAGCCCATCTTAAGGCACCTTGTGGAAGG	180			
Qy	283	CAGAATGCCAGTGTGCTTATGGACCCACAGGAGCTGGGAAGACGCACCAATGCTG	342			
Db	181	CAGAATGCCAGTGTGCTTATGGACCCACAGGAGCTGGGAAGACGCACCAATGCTG	240			
Qy	343	GGCAGCCAGAGCAACCTGGGGTATCCCGGGCTCTCATGACCTCTCTGACGCTCACA	402			
Db	241	GGCAGCCAGAGCAACCTGGGGTATCCCGGGCTCTCATGACCTCTCTGACGCTCACA	300			
Qy	403	AGGAGGAGGTGCGGAGGCGGCCCATGCGGCTTCTGTACCACTGCTTACCTAGAG	462			
Db	301	AGGAGGAGGTGCGGAGGCGGCCCATGCGGCTTCTGTACCACTGCTTACCTAGAG	360			
Qy	463	ATCTACAGGAGAGGTATTAGACCTCTCTGGAACCTGCTTCTGGGAGACCTGTAATCCGA	522			
Db	361	ATCTACAGGAGAGGTATTAGACCTCTCTGGAACCTGCTTCTGGGAGACCTGTAATCCGA	420			
Qy	523	GAAGACTGCGGGGAATATCTGATTCGGGTCTCTCCAGAGCCATCAGTAGCTTT	582			
Db	421	GAAGACTGCGGGGAATATCTGATTCGGGTCTCTCCAGAGCCATCAGTAGCTTT	480			
Qy	583	GCTGATTTGAGCGGCACTTCTCCAGCCAGTGCAGAAATCGAGCTGTAGAGCAACCCG	642			
Db	481	GCTGATTTGAGCGGCACTTCTCCAGCCAGTGCAGAAATCGAGCTGTAGAGCAACCCG	540			
Qy	643	CTCAACCAAGCCTCTCCCGAGTATGCTGTGCTCTGTCATAGGTGGACCCAGCGGA	702			
Db	541	CTCAACCAAGCCTCTCCCGAGTATGCTGTGCTCTGTCATAGGTGGACCCAGCGGA	600			
Qy	703	CGTTTGGCCCATTTGCGCCAGGAGGAGAAATCTACCTGATTTGCTGGCTGGGTCA	762			
Db	601	CGTTTGGCCCATTTGCGCCAGGAGGAGAAATCTACCTGATTTGCTGGCTGGGTCA	660			
Qy	763	GAGGACAAACCGCGCACAGGCAACAGGGCTTCCGCTTAAAGAGAGTGAGGACCATCAAC	822			
Db	661	GAGGACAAACCGCGCACAGGCAACAGGGCTTCCGCTTAAAGAGAGTGAGGACCATCAAC	720			
Qy	823	ACCTCCCTGTTTCTGCGCAAGTGTAGTGGCTGATCAGGCTCCCTCGTGTA	882			
Db	721	ACCTCCCTGTTTCTGCGCAAGTGTAGTGGCTGATCAGGCTCCCTCGTGTA	780			
Qy	883	CCCTTATCGGACAGCAAGCTCACTCCCTATTGACGAGACTCTCTGGGTGGCTCAGCCAC	942			
Db	781	CCCTTATCGGACAGCAAGCTCACTCCCTATTGACGAGACTCTCTGGGTGGCTCAGCCAC	840			
Qy	943	AGTATCCTTATTCGCAATGCGCTCTGAGAGACGCTTCTACCTAGACAGAGTCTCCGA	1002			
Db	841	AGTATCCTTATTCGCAATGCGCTCTGAGAGACGCTTCTACCTAGACAGAGTCTCCGA	900			

Qy	1003	CTCAACTTTCTGCCAGG-TCCAAGGAGGTGATCAATCGGCCTTT	1047			
Db	901	CTCAACTTTCTGCCAGGTTCCCAAGGAGGTGATCCATCGGCCTTT	946			
RESULT 10		AL559772	1083 bp	mrna	linear	EST 02-APR-2004
LOCUS		AL559772	Homo sapiens B CELLS (RAMOS CELL LINE)	Homo sapiens	CDNA	
DEFINITION		clone CS0DG005Y111 5-PRIME, mRNA sequence.				
ACCESSION		AL559772				
VERSION		AL559772.3	GI:46185144			
KEYWORDS		EST.				
SOURCE		Homo sapiens (human)				
ORGANISM		Homo sapiens				
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE		1 (Bases 1 to 1083)				
AUTHORS		Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
TITLE		Full-length cDNA libraries and normalization				
JOURNAL		Unpublished (2001)				
COMMENT		On Feb 15, 2001 this sequence version replaced gi:31283903. Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7766.r For more information about this cluster, see http://www.genoscope.cns.fr/cdna?c=CS0DG005CF06QP1&c=7766.r.				
FEATURES		Location/Qualifiers				
		1..1083				
		/organism="Homo sapiens"				
		/mol_type="mrna"				
		/db_xref="taxon:9606"				
		/clone="CS0DG005Y111"				
		/tissue_type="B CELLS (RAMOS CELL LINE)"				
		/cell_line="RAMOS CELLS LINE"				
		/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"				
		/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."				
ORIGIN		Query Match	62.3%;	Score 912.8;	DB 1;	Length 1083;
		Best Local Similarity	96.1%;	Pred. No. 2.5e-232;		
		Matches	942;	Conservative	13;	Mismatches 21; Indels 4; Gaps 2;
Qy	4	GGTGCTGTGCGCTAAGCAAGATTGGAGCTACTGCTGCTCCACCTCGCTAAGG	63			
Db	69	GGTGGCTGTGCGCTAAGCAAGATTGGAGCTACTGCTGCTCCACCTCGCTAAGG	128			
Qy	64	GTGCTGTGCGCTGCGGCTATTTGGATGGAACAGCGGGAGCAAGTATCCCCCTGT	123			
Db	129	GTGCTGTGCGCTGCGGCTATTTGGATGGAACAGCGGGAGCAAGTATCCCCCTGT	188			
Qy	124	GTGCGGGCATGACAGCTGCTCTTAGAGATTGCTAAGTGGAGAACCCACGAGGACT	193			
Db	189	GTGCGGGCATGACAGCTGCTCTTAGAGATTGCTAAGTGGAGAACCCACGAGGACT	248			
Qy	184	CTCAATACAGTTTGTATGCTTCTATGGGAGGAGTACTCAGCAGGACATCTATGCA	243			
Db	249	CTCAATACAGTTTGTATGCTTCTATGGGAGGAGTACTCAGCAGGACATCTATGCA	308			
Qy	244	GGTTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	303			

Db 309 GGTTCAGTGCAGCCATCCTTAAGGACCTTGTCTGGAGGGCAGATGCCAGTGTGCTTGCC 368
Qy 304 TATGGACCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGCCACAGAGCAACCTGGG 363
Db 369 TATGGACCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGCCACAGAGCAACCTGGG 428
Qy 364 GTGATCCCGGGCTCTCATGACCTCTGAGCTTCAAGAGGAGGAGGTGCGGAGGGC 423
Db 429 GTGATCCCGGGCTCTCATGACCTCTGAGCTTCAAGAGGAGGAGGTGCGGAGGGC 488
Qy 424 CGGCATCGGGCCCTTCTGTCACCATGCTTACCTAGAGATCTACACAGGAGAAGTATTA 483
Db 489 CGGCATCGGGCCCTTCTGTCACCATGCTTACCTAGAGATCTACACAGGAGAAGTATTA 548
Qy 484 GACCTCTGGACCCCTTCTGGGAGACCTGGTAAATCCGAGAGACTGCGGGGGAATATC 543
Db 549 GACCTCTGGACCCCTTCTGGGAGACCTGGTAAATCCGAGAGACTGCGGGGGAATATC 608
Qy 544 CTGATTCGGGTCTCTCCAGAGCCATCATGAGTCTGATTTGATTTGAGCGGCACTTC 603
Db 609 CTGATTCGGGTCTCTCCAGAGCCATCATGAGTCTGATTTGATTTGAGCGGCACTTC 668
Qy 604 CTGCAGCCAGTCGGAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCCCGC 663
Db 669 CTGCAGCCAGTCGGAATCGGACTGTAGGAGCCACCGGCTCAAMWAGCGCTCTCCCGC 728
Qy 664 AGTCATGCTGTCTCTGCTCAAGTGGACGAGGGAACCTTTGGGCCCAATTTTCGCAG 723
Db 729 AGTCATGCTGTCTCTGCTCAAGTGGACGAGGGAACCTTTGGGCCCAATTTTCGCAG 788
Qy 724 CGAGAGGGAATCTCTACCTGATTTGACTTGGCTGGTTCAGAGGCAACCGCGGCACAGGC 783
Db 789 CGAGAGGGAATCTCTACCTGATTTGACTTGGCTGGTTCAGAGGCAACCGCGGCACAGGC 848
Qy 784 AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCCTGTTGTCTCGGC 843
Db 849 AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCCTGTTGTCTCGGC 908
Qy 844 AAAGTGTAGATGCGCTGAATCAGGGCTCCCTCGTGTACCTTATCGGAGCAGCAAGCTC 903
Db 909 AAAGTGTAGATGCGCTGAATCAGGGCTCCCTCGTGTACCTTATCGGAGCAGCAAGCTC 968
Qy 904 ACTCGCTATTCAGGAGCTCTCTGGTGGCTCAGCCACAGTATCTTTATTCGCAACATT 963
Db 969 CTYCGCTATTCGA-GRCTCTCTGGTGGTCAACCCAGTAT---CVTATTGCAAAATTG 1024
Qy 964 GCCCTCAGAGACGCTTCTA 983
Db 1025 CCCTKARAGCTTTTHCTA 1044

RESULT 11
AL518361 1016 bp mRNA linear EST 23-MAR-2004
LOCUS
DEFINITION AL518361 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
CSODA009YD01 5-PRIME, mRNA sequence.
ACCESSION AL518361
VERSION AL518361.3 GI:45654904
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1016)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 13, 2001 this sequence version replaced gi:30536600.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.
This sequence belongs to sequence cluster 7766.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CSODA009CB01QPI&c=7766.r.
Location/Qualifiers
1. 1016
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODA009YD01"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

FEATURES

source

ORIGIN

Query Match 61.3%; Score 897; DB 1; Length 1016;
Best Local Similarity 98.5%; Pred. No. 4.1e-228;
Matches 923; Conservative 3; Mismatches 8; Indels 3; Gaps 2;
Qy 4 GGTGCTGTGGCTTAAGCAAGATTGGAGCTACTCGTCCACCTCCAGCTCGGTAAGG 63
Db 9 GGTGCTGTGGCTTAAGCAAGATTGGAGCTACTCGTCCACCTCCAGCTCGGTAAGG 68
Qy 64 GTGCTGTGGCTGTGGCCATTTGTGGATGGAAACAGCGGAGCAAGTGATCCCCCTGT 123
Db 69 GTGCTGTGGCTGTGGCCATTTGTGGATGGAAACAGCGGAGCAAGTGATCCCCCTGT 128
Qy 124 GTGCGGGCATGGACAGCTGCTCTAGAGATTCTAACTGGAGAACCAACAGAGACT 183
Db 129 GTGCGGGCATGGACAGCTGCTCTAGAGATTCTAACTGGAGAACCAACAGAGACT 188
Qy 184 CTCAATACCAAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAGGAGCATCTATGCA 243
Db 189 CTCAATACCAAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAGGAGCATCTATGCA 248
Qy 244 GGTTCAGTCAGCCCATCTTAAGGCACTTGTGGAAGGGCAGAAATGCCAGTGTGCTTGGC 303
Db 249 GGTTCAGTCAGCCCATCTTAAGGCACTTGTGGAAGGGCAGAAATGCCAGTGTGCTTGGC 308
Qy 304 TATGGACCCACAGGAGCTGGGAGACGCAACAATGTGGGAGCCAGAGCAACCTGGG 363
Db 309 TATGGACCCACAGGAGCTGGGAGACGCAACAATGTGGGAGCCAGAGCAACCTGGG 368
Qy 364 GTGATCCCGGGCTCTCATGGACCTCCTGCAGCTCAACAGGAGGAGGTGCGGAGGGC 423
Db 369 GTGATCCCGGGCTCTCATGGACCTCCTGCAGCTCAACAGGAGGAGGTGCGGAGGGC 428
Qy 424 CGGCCATGGGCCCTTTCTGTCAACATGCTTACCTAGAGATCTACACAGGAGAAGTATTA 483
Db 429 CGGCCATGGGCCCTTTCTGTCAACATGCTTACCTAGAGATCTACACAGGAGAAGTATTA 488
Qy 484 GACCTCTGGACCCCTTCTGTCGAGAGCTGGTAAATCCGAGAGACTGCGGGGGAATATC 543
Db 489 GACCTCTGGACCCCTTCTGTCGAGAGCTGGTAAATCCGAGAGACTGCGGGGGAATATC 548
Qy 544 CTGATTCGGGTCTCTCCAGAGCCATCAGTAGCTTTGATTTGAGCGGCACTTC 603
Db 549 CTGATTCGGGTCTCTCCAGAGCCATCAGTAGCTTTGATTTGAGCGGCACTTC 608
Qy 604 CTGCCAGCCAGTCGGAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCCCGC 663
Db 609 CTGCCAGCCAGTCGGAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCCCGC 668
Qy 664 AGTCATGCTGTCTCTGCTCAAGTGGACGAGGGAACCTTTGGGCCCAATTTTCGCAG 723
Db 669 AGTCATGCTGTCTCTGCTCAAGTGGACGAGGGAACCTTTGGGCCCAATTTTCGCAG 728

[illegible][illegible]

Query Match	59.8%; Score 875.4; DB 5; Length 971;
Best Local Similarity	97.2%; Pred. No. 2.4e-222;
Matches 934; Conservative	0; Mismatches 21; Indels 6; Gaps 4;
QY	188 AATACCAGTTTCATGCTTCTATGGGAGAGAGTAGTACTCAGCAGGACATCTATGCAAGTT 247
DB	
QY	2 AATACAGTTTATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCAAGTT 61
DB	
QY	248 CAGTCAGGCCATCTTAAGGCATCTTGCTGGAAGGGCAGAAATGCCAGTGTGCTTGCCATG 307
DB	
QY	62 CAGTCAGGCCATCTTAAGGCATCTTGCTGGAAGGGCAGAAATGCCAGTGTGCTTGCCATG 121
DB	
QY	308 GACCCACAGAGCTGGGAGAGCGCACAAATGCTGGGAGCCACAGAGCAACTTGGGGTCA 367
DB	
QY	122 GACCCACAGAGCTGGGAGAGCGCACAAATGCTGGGAGCCACAGAGCAACTTGGGGTCA 181
DB	
QY	368 TCCCGGGGCTCTCATGSGACCTCCTCAGCTCAACAGGGAGGAGGTGCGAGGGCCGCG 427
DB	
QY	182 TCCCGGGGCTCTCATGSGACCTCCTCAGCTCAACAGGGAGGAGGTGCGAGGGCCGCG 241
DB	
QY	428 CATGGGCCCTTCTGTTCACATGTCTTACCTAGAGATCTACAGGAGAGAGTATTAGACC 487
DB	
QY	242 CATGGGCCCTTCTGTTCACCAATGCTTACCTAGAGATCTACAGGAGAGAGTATTAGACC 301
DB	
QY	488 TCTCTGACCCCTGCTTCGGGAGACCTGGTAACTCGAGAGACTCCCGGGGGAATATCCTGA 547
DB	
QY	302 TCTCTGACCCCTGCTTCGGGAGACCTGGTAACTCGAGAGACTCCCGGGGGAATATCCTGA 361
DB	
QY	548 TTCCGGGTCTCTCCCAAGAACCCATCAGTAGCTTTGTGATTTTGAACGGGCACTTCTCTGC 607
DB	
QY	362 TTCCGGGTCTCTCCCAAGAACCCATCAGTAGCTTTGTGATTTTGAACGGGCACTTCTCTGC 421
DB	
QY	608 CAGCCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACAGAGCGCTCTCCCGCAGTC 667
DB	
QY	422 CAGCCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACAGAGCGCTCTCCCGCAGTC 481
DB	
QY	668 ATGCTGTGCTCTCTGAGTGAAGTGGACAGCGGGAAACGTTTGGGCCCATTTTCGCCAGCGAG 727
DB	
QY	482 ATGCTGTGCTCTCTGAGTGAAGTGGACAGCGGGAAACGTTTGGGCCCATTTTCGCCAGCGAG 541
DB	
QY	728 AGGGAAACTCTACCTGAATTGACTTGCTGGGTGAGAGCAACCCGGCGCAGAGGCAACA 787
DB	
QY	542 AGGGAAACTCTACCTGAATTGACTTGCTGGGTGAGAGCAACCCGGCGCAGAGGCAACA 601
DB	
QY	788 AGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCCCTGTTTGTCTCGGCCAAAG 847
DB	
QY	602 AGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCCCTGTTTGTCTCGGCCAAAG 661
DB	
QY	848 TGGTAGATGGGCTGAATCAGGGCTCTCCTGTGTACCTTATCGGACAGCAAGTCACTC 907
DB	
QY	662 TGGTAGATGGGCTGAATCAGGGCTCTCCTGTGTGTACCTTATCGGACAGCAAGTCACTC 721
DB	
QY	908 GCCTATTGCAAGACTCTCTGGTGGCTCAGCCCAAGTATCTTATTGGCAACATTGCCCC 967
DB	
QY	722 GCCTATTGCAAGACTCTCTGGTGGCTCAGCCCAAGTATCTTATTGGCAACATTGCCCC 781
DB	
QY	968 CTGAGAGAGCGTCTTACTAGACACAGTCTCCGACCTCAACTTTGCTG-CCAGTCCCAAG 1026
DB	
QY	782 CTGAGAGAGCGTCTTACTAGACACAGTCTCCGACCTCAACTTTGCTG-CCAGTCCCAAG 841
DB	
QY	1027 GAGGTGATCAATCGGCCCTTTTACCAATGAGAGCTGACGCTCATGCTGTT-GGGACCTGT 1085
DB	
QY	842 GAGGTGAATCATCGGGCTTTTACCAATGAGAGCTGACGCTCATGCTGTTGCTGGGACCTGT 901
DB	
QY	1086 TAAGCTGTC--TCAGAGAGATTGCTT--GGTCCACAGAGGCAAGAGAGCCCGCAGGCC 1141
DB	
QY	902 TAAGCTGCGCTCAAAAAGGAATGGCTTGGGGCCACCCCAAGGCCAAAGAGAGCCCGCAGGCC 961
DB	
QY	1142 C 1142

3

[illegible]

Db	373	TATGGACCCACAGAGCTGGGAACGACACACATGCTGGGCAGCCCGAGAGCAACCTGGG	432
Qy	364	GTGATCCCGGGGCTCTCATGACCTCTGACGCTCAAGGGAGGAGGTGCCGAGGGC	423
Db	433	GTGATCCCGGGGCTCTCATGACCTCTGACGCTCAAGGGAGGAGGTGCCGAGGGC	492
Qy	424	CGGCCATGGGCCCTTTCTGTCCACCATGTCTTACCTAGAGATCTACAGGAGAGGTATTA	483
Db	493	CGGCCATGGGCCCTTTCTGTCCACCATGTCTTACCTAGAGATCTACAGGAGAGGTATTA	552
Qy	484	GACCTCTGGACCTGCTTGGGAGACTGTGTAATCCGAGAGACTGCCGGGGGAATATC	543
Db	553	GACCTCTGGACCTGCTTGGGAGACTGTGTAATCCGAGAGACTGCCGGGGGAATATC	612
Qy	544	CTGATTCGGGTCTCTCCAGAACCCCATCAGTAGCTTTGCTGATTTTGAGCGGCATTTC	603
Db	613	CTGATTCGGGTCTCTCCAGAACCCCATCAGTAGCTTTGCTGATTTTGAGCGGCATTTC	672
Qy	604	CTGCCAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGGGCTCAACCGAGCGTCTCTCCGC	663
Db	673	CTGCCAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGGGCTCAACCGAGCGTCTCTCCGC	732
Qy	664	AGTCATGCTGTGCTCTGCTCAAGGTGAGCGAGCGGGAACGTTTGGCCCCCATTTCCGCCAG	723
Db	733	AGTCATGCTGTGCTCTGCTCAAGGTGAGCGGGAACGTTTGGCCCCCATTTCCGCCAG	792
Qy	724	CGAGAGGGAAAACTCTACCTGATTGACTTGGCTGGGTGAGAGGACAAACCGCGCACAGGC	783
Db	793	CGAGAGGGAAAACTCTACCTGATTGACTTGGCTGGGTGAGAGGACAAACCGCGCACAGGC	852
Qy	784	AACAGGGCCCTTCGGCTTAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTGGGC	843
Db	853	AACAAGGGCCCTTCGGCTTAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTGGGC	912
Qy	844	-AAAGTGGTAGATGCGCTGAATCAGGGCCCTCCCTCGTGTACCTTAT	888
Db	913	AAAGTGGTAGATGCGCTGAATCAGGGCCCTCCCTCGTGTACCTTAT	958

Search completed: November 10, 2004, 18:18:12
 Job time : 5093.37 secs

This Page Blank (uspto)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	1115	100.0	1115	6	AR210052	Sequence
2	1112.4	99.8	1538	6	AR210053	Sequence
3	1107.4	99.3	1398	9	BT007259	Homo sapi
4	1107.4	99.3	1398	12	BT007888	Synthetic
5	1107.4	99.3	2097	6	AR304057	Sequence
6	1107.4	99.3	2097	9	AB017430	Homo sapi
7	1107.4	99.3	2117	9	BC028155	Homo sapi
8	1107.4	99.3	2134	9	BC004352	Homo sapi
9	1038	93.1	1041	6	AR210054	Sequence
10	1035.4	92.3	1464	6	AR210055	Sequence
11	779.8	69.9	2086	10	BC003427	Mus muscu
12	681.4	61.1	151041	2	AC101919	Mus muscu
13	681.4	61.1	164759	10	AC101752	Mus muscu
14	654.2	58.7	297639	2	AC101908	Mus muscu
15	522.8	46.9	196674	10	AC102127	Mus muscu
16	462.2	41.5	2756	5	BC063896	Xenopus t
17	454.6	40.8	2565	5	BC043733	Xenopus l
18	454.6	40.8	2583	5	XL249841	Xenopus l
19	454.6	40.8	2640	5	BC073177	Xenopus l

Db	241	GGAAACCA	CCAGGAGACTCTCAAA	TACAGTTTGATGCTTCTATGGGGAGAGGACTC	300
Qy	301	AGCAGGACATCTATG	CAGGTTTCAGTGCAGGCCATCTTAAGGCATCTGCTGGAAGGGCAGA	360	
Db	301	AGCAGGACATCTATG	CAGGTTTCAGTGCAGGCCATCTTAAGGCATCTGCTGGAAGGGCAGA	360	
Qy	361	ATGCCAGTGCTCTTGC	CTATGGACCCA	CAGGAGCTGGGAAAGACACAAATGCTGGGCA	420
Db	361	ATGCCAGTGCTCTTGC	CTATGGACCCA	CAGGAGCTGGGAAAGACACAAATGCTGGGCA	420
Qy	421	GCCCAGAGCAACTCTG	GGGTCATCCCGGGGCTCTCATGGACCTCCTGCAGCTCACAAAGG	480	
Db	421	GCCCAGAGCAACTCTG	GGGTCATCCCGGGGCTCTCATGGACCTCCTGCAGCTCACAAAGG	480	
Qy	481	AGGAGGTGCCGAGGG	CGGGCCATGGGCCCTTTCTGTCAACATGTCTTACTTAGAGATCT	540	
Db	481	AGGAGGTGCCGAGGG	CGGGCCATGGGCCCTTTCTGTCAACATGTCTTACTTAGAGATCT	540	
Qy	541	ACAGGAGAGGTTATG	ACCTCTGTGACCCCTGCTTCGGGAGACCTGGTAAATCCGAGAAG	600	
Db	541	ACAGGAGAGGTTATG	ACCTCTGTGACCCCTGCTTCGGGAGACCTGGTAAATCCGAGAAG	600	
Qy	601	ACTGCCGGGGGAATAT	CTGATTCGGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGGCTG	660	
Db	601	ACTGCCGGGGGAATAT	CTGATTCGGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGGCTG	660	
Qy	661	ATTTTAGAGGGCACTT	CTGCTGCAGCCAGTGGAAATCGGACTGTAGGAGCCACCCGGCTCA	720	
Db	661	ATTTTAGAGGGCACTT	CTGCTGCAGCCAGTGGAAATCGGACTGTAGGAGCCACCCGGCTCA	720	
Qy	721	ACAGCGCTCTCTCCG	CAGTCATGCTGTCTCTGGTCAAGSTGGACCGCGGGAAACGTT	780	
Db	721	ACAGCGCTCTCTCCG	CAGTCATGCTGTCTCTGGTCAAGSTGGACCGCGGGAAACGTT	780	
Qy	781	TGGCCCAATTTGCG	CAGGAGGAAAACTCTACCTGATTGACTTGGCTGGGTCAGAGG	840	
Db	781	TGGCCCAATTTGCG	CAGGAGGAAAACTCTACCTGATTGACTTGGCTGGGTCAGAGG	840	
Qy	841	ACAAACGGCGCACAGG	CAAAAGGGCCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCT	900	
Db	841	ACAAACGGCGCACAGG	CAAAAGGGCCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCT	900	
Qy	901	CCCTGTTTGCTCTGGG	CAAAAGTGTAGATGCGCTGAAATCAGGGGCTCCCTCGTGTACCTT	960	
Db	901	CCCTGTTTGCTCTGGG	CAAAAGTGTAGATGCGCTGAAATCAGGGGCTCCCTCGTGTACCTT	960	
Qy	961	ATCGGGA	CAGCAAGCTCACTCGCCTATTGAGGACTCTCTGGGTGGCTCAGCCCCACAGTA	1020	
Db	961	ATCGGGA	CAGCAAGCTCACTCGCCTATTGAGGACTCTCTGGGTGGCTCAGCCCCACAGTA	1020	
Qy	1021	TCCTTATTGCCAAAT	TGCCCCCTCAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCA	1080	
Db	1021	TCCTTATTGCCAAAT	TGCCCCCTCAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCA	1080	
Qy	1081	ACTTTGCTGCCAGT	CCAAAGGAGTGATCAATTGA	1115	
Db	1081	ACTTTGCTGCCAGT	CCAAAGGAGTGATCAATTGA	1115	

FEATURES	source	Location/Qualifiers
	1. .1538	
	/organism="unknown"	
	/mol_type="unassigned DNA"	
ORIGIN		
Query Match	99.8%;	Score 1112.4; DB 6; Length 1538;
Best Local Similarity	99.9%;	Pred. No. 1.3e-277;
Matches 1113; Conservative	0; Mismatches	1; Indels 0; Gaps 0;
Qy	1	ATGCAGCGCGGGCGCTCGACGACGAGAGGGCGACGCGAGATGGCGGACAGCTTCACGGG 60
Db	1	ATGCAGCGCGGGCGGCTCGACGACGAGAGGGCGACGCGAGATGGCGGACAGCTTCACGGG 60
Qy	61	CGGCGATCTCAGAGAGCTGGTCGGTCTGCGTAAAGCAAGATTGGAGCTATCTGTCGTCCAC 120
Db	61	CGGCGATCTCAGAGAGCTGGTCGGTCTGCGTAAAGCAAGATTGGAGCTATCTGTCGTCCAC 120
Qy	121	CTCCAGCTCCGGTAAGGGTGGCTGTGGACTGCGGCCATTGTGGATGGAAACAGCGGGAG 180
Db	121	CTCCAGCTCCGGTAAGGGTGGCTGTGGACTGCGGCCATTGTGGATGGAAACAGCGGGAG 180
Qy	181	CAAGTGATCCCCCTGTGTGGGGGACATGACAGCTGCTCTCTAGAGATTGCTTAACCTGGA 240
Db	181	CAAGTGATCCCCCTGTGTGGGGGACATGACAGCTGCTCTCTAGAGATTGCTTAACCTGGA 240
Qy	241	GGAAACACAGGAGACTCTCAATACACAGTTTGATGCTTCATATGGGAGAGGAGTACTC 300
Db	241	GGAAACACAGGAGACTCTCAATACACAGTTTGATGCTTCATATGGGAGAGGAGTACTC 300
Qy	301	AGCAGGACATCTATGCAAGTTCAGTGCAGGCCATCTCTAAGGCATCTTCTCGAAGGGCAGA 360
Db	301	AGCAGGACATCTATGCAAGTTCAGTGCAGGCCATCTCTAAGGCATCTTCTGGAAGGGCAGA 360
Qy	361	ATGCCAGTGTCTTGGCCCTATGGAACCCACAGAGCTGGGAAGACGCAACAATGCTGGGCA 420
Db	361	ATGCCAGTGTCTTGGCCCTATGGAACCCACAGAGCTGGGAAGACGCAACAATGCTGGGCA 420
Qy	421	GCCCAGAGCAACCTGGGGTGATCCCGGGGCTCTCATGAGCACTCTCGAGCTCACAAGGG 480
Db	421	GCCCAGAGCAACCTGGGGTGATCCCGGGGCTCTCATGAGCACTCTCGAGCTCACAAGGG 480
Qy	481	AGGAGGTCGCGAGGGCGGCCATATGGGCCCTTTTGTCACCATGTCTTAAGTAGAGATCT 540
Db	481	AGGAGGTCGCGAGGGCGGCCATATGGGCCCTTTTGTCACCATGTCTTAAGTAGAGATCT 540
Qy	541	ACGAGGAGAGGTAATTAGACTCTCTGACACCTCTCGGAGAGCTCGTAAATCCGAGAAG 600
Db	541	ACGAGGAGAGGTAATTAGACTCTCTGACACCTCTCGGAGAGCTCGTAAATCCGAGAAG 600
Qy	601	ACTGCCGGGGAAATATCTCTGATTCGGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTG 660
Db	601	ACTGCCGGGGAAATATCTCTGATTCGGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTG 660
Qy	661	ATTTTGAGCGGCACTTCCTGCGACGCAAGTGGAAATCGGACTGTAGGAGCCACCCGGCTCA 720
Db	661	ATTTTGAGCGGCACTTCCTGCGACGCAAGTGGAAATCGGACTGTAGGAGCCACCCGGCTCA 720
Qy	721	ACGAGCGTCTCTCCGCGAGTCATGCTGTGCTCTGCTCTGCTCAAGGTGGACGAGCGGAAACGTT 780
Db	721	ACGAGCGTCTCTCCGCGAGTCATGCTGTGCTCTGCTCTGCTCAAGGTGGACGAGCGGAAACGTT 780
Qy	781	TGGGCCCATTTTCGACGAGAGGAAACCTCTACCTGATTGACTTGGCTGGGTACAGAGG 840
Db	781	TGGGCCCATTTTCGACGAGAGGAAACCTCTACCTGATTGACTTGGCTGGGTACAGAGG 840
Qy	841	ACAACCGCGCACAGGCAACAAGGGCTTCGGCTTAAAGAGATGGAGCCATCAACACCT 900
Db	841	ACAACCGCGCACAGGCAACAAGGGCTTCGGCTTAAAGAGATGGAGCCATCAACACCT 900
Qy	901	CCCTGTTTGTCTCGGCAAGTGTAGATGCGCTGAATCAGGGGCTCCCTCGTGTACCTT 960
Db	901	CCCTGTTTGTCTCGGCAAGTGTAGATGCGCTGAATCAGGGGCTCCCTCGTGTACCTT 960

QY 961 ATCGGACAGCAAGCTCACTCGCTATTGAGGACTCTCTGGTGGCTCAGGCCACAGTA 1020
 DB ATCGGACAGCAAGCTCACTCGCTATTGAGGACTCTCTGGTGGCTCAGGCCACAGTA 1020
 QY 1021 TCCATTATGCCAATGTCCTGAGAGAGCTTCTACCTAGACAGCTCCTCGCACTCA 1080
 DB TCCATTATGCCAATGTCCTGAGAGAGCTTCTACCTAGACAGCTCCTCGCACTCA 1080
 QY 1081 ACTTCTGCTGAGGTCAGGAGGCTGATCAATG 1114
 DB ACTTCTGCTGAGGTCAGGAGGCTGATCAATG 1114

RESULT 3
 LOCUS BT007259 1998 bp mRNA linear PRI 13-MAY-2003
 DEFINITION Homo sapiens kinesin-like 4 mRNA, complete cds.
 ACCESSION BT007259
 VERSION BT007259.1 GI:30583356
 KEYWORDS FLI CDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1998)
 Kalline, N., Chen, X., Rolfs, A., Halleck, A., Hines, L., Eisenstein, S.,
 Koundinya, M., Raphael, J., Moreira, D., Kelley, T., LaBaer, J., Lin, Y.,
 Phelan, M. and Farmer, A.
 Cloning of human full-length cDNAs in BD Creator(TM) System Donor
 vector

Unpublished
 2 (bases 1 to 1998)
 Kalline, N., Chen, X., Rolfs, A., Halleck, A., Hines, L., Eisenstein, S.,
 Koundinya, M., Raphael, J., Moreira, D., Kelley, T., LaBaer, J., Lin, Y.,
 Phelan, M. and Farmer, A.
 Direct Submission
 Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow
 Circle, Palo Alto, CA 94303, USA

This cDNA clone is a part of a collection of human full length
 expression clones generated by BD Biosciences Clontech and the
 Harvard Institute of Proteomics. Each cDNA has been cloned in two
 forms: with and without stop-codon (to allow fusion with C-terminal
 tag). The cDNA has been directionally cloned using BD In-Fusion(TM)
 cloning system between the Sali and HindIII sites of the pDNR-DUAL
 vector. Additional sequences in the clone: 'ACC' after Sali site
 and before 'ATG' to provide Kozak consensus sequence; 'GG' after
 last codon and before HindIII site to maintain reading frame.
 Clone distribution: <http://bioinfo.clontech.com/orfclones>.

Location/Qualifiers
 1..1998
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="GH00245X1.0"
 /clone_lib="BD Creator(TM) CDS Library derived from MGC
 collection"
 /lab_host="DH5alpha Ti resistant"
 /note="Vector: pDNR-Dual"
 1..1998
 /codon_start=1
 /product="kinesin-like 4"
 /protein_id="AAP35923.1"
 /db_xref="GI:30583357"

/translation="MAAGSTQQRREMAAASAAISGAGCRISKIGATRRPPRV
 RVNLRPFVDTAGSDPPCVGRMDSCSLEIANWRNHQETLKQFDIFYGERTQD
 IYAGSVQPIRLHLLGQNASVLYAGTGTAGKTHMLGSPQPGVPIPRALMDLQITRE
 EGAERPNALSVTMSYLBIEYOKVLDLPASGDLVIREDCRGNLIPGLSKPTSSP
 ADPERHLPASNRVTGATRLNQRSSSHAVLLVKDQRELRAPRQEGKLYLIDLA
 GSDNRNRTGNKRLKESGINTSLFVGVKVDALNQGLRPVPRDRLSKLRLQDSLG
 GSASHILIANAPERFVLDYTSALNFAARKEVINRPTNIESLQHPALGPVKSQSE
 LLGPEAKRARGPEEEIIGSPMAAPASQKLSPLQKLSMDPAMLERLLSLDRLL
 ASQSQGAPLLSTPKRERWLMKTVEEKDLETERLKTKQKELEAKMLQAQBAKENHC

PTMLRPLSHRTVTGAKPLKKA VVMPQLIQEOAASPNAIEHILKNKRKRKLKESLDAL
 EBEBAEDCWELQISPPELLAHGRKILDLNFSARDLRSIORIGPKQAQLVGVREL
 HGPPFOVEDLERVEGITGQMESFLKANILGLAAGORCCAS"
 ORIGIN
 Query Match 99.3%; Score 1107.4; DB 9; Length 1998;
 Best Local Similarity 99.9%; Pred. No. 2.6e-276;
 Matches 1108; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 6 GCGCGGGGGGCTCGACGAGGAGGCGACCGAGATGGCGGACGCTTCAGCGGGGGG 65
 DB 4 GCGCGGGGGGCTCGACGAGGAGGCGACCGAGATGGCGGACGCTTCAGCGGGGGG 63
 QY 66 ATCTCAGGAGCTGGTGGCTTAAGCAAGATTGGAGCTACTCTGTCACCTCCA 125
 DB 64 ATCTCAGGAGCTGGTGGCTTAAGCAAGATTGGAGCTACTCTGTCACCTCCA 123
 QY 126 GCTCGCGTAAGGGTGGCTGTCGACTGCGGCGCATTTGTGGATGGAAACAGCGGGAGCAAGT 185
 DB 124 GCTCGCGTAAGGGTGGCTGTCGACTGCGGCGCATTTGTGGATGGAAACAGCGGGAGCAAGT 183
 QY 186 GATCCCCCTGTCGGGGGCGATGCAAGCTGCTCTTAGAGATTGCTAACTGGAGGAAC 245
 DB 184 GATCCCCCTGTCGGGGGCGATGCAAGCTGCTCTTAGAGATTGCTAACTGGAGGAAC 243
 QY 246 CACGAGGAGACTCTCAAAATACCAATTGATGCTTCTATGGGGAGAGAGTAGTACTCAGCAG 305
 DB 244 CACGAGGAGACTCTCAAAATACCAATTGATGCTTCTATGGGGAGAGAGTAGTACTCAGCAG 303
 QY 306 GACATCTATGCAAGTTTCAGTGCAGGCCATCTTAAGGCACTTCTGCTGGAGGGAGAGATGCC 365
 DB 304 GACATCTATGCAAGTTTCAGTGCAGGCCATCTTAAGGCACTTCTGCTGGAGGGAGAGATGCC 363
 QY 366 AGTGTGCTTGCCTATGGACCCACAGGAGCTGGGAAGAGACACAAATGCTGGGCGGCCCA 425
 DB 364 AGTGTGCTTGCCTATGGACCCACAGGAGCTGGGAAGAGACACAAATGCTGGGCGGCCCA 423
 QY 426 GAGCAACCTGGGTGATCCCGGGGCTCTCATGGACCTCTCGACCTCAAGAGGAGGAG 485
 DB 424 GAGCAACCTGGGTGATCCCGGGGCTCTCATGGACCTCTCGACCTCAAGAGGAGGAG 483
 QY 486 GGTCCGAGGGCGGCCATGGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTACACAG 545
 DB 484 GGTCCGAGGGCGGCCATGGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTACACAG 543
 QY 546 GAGAAGGTATTAGACTCTCTGACCCCTGTCGGAGAGCTGGTAAATCGGAGAAAGCTGC 605
 DB 544 GAGAAGGTATTAGACTCTCTGACCCCTGTCGGAGAGCTGGTAAATCGGAGAAAGCTGC 603
 QY 606 CGGGGGAATATCTGATTCCGGGCTCTCCAGAGAGCCCATCAGTAGCTTTTCTGATTTT 665
 DB 604 CGGGGGAATATCTGATTCCGGGCTCTCCAGAGAGCCCATCAGTAGCTTTTCTGATTTT 663
 QY 666 GAGCGGCACTTCTCGCCAGCCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACCCAG 725
 DB 664 GAGCGGCACTTCTCGCCAGCCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACCCAG 723
 QY 726 CGTCTCTCCGAGTCATGCTGCTCTCTGTCCTGAGGTGGACCCAGCGGGAGAGCTTTGGCC 785
 DB 724 CGTCTCTCCGAGTCATGCTGCTCTCTGTCCTGAGGTGGACCCAGCGGGAGAGCTTTGGCC 783
 QY 786 CCATTTCCCGAGCGAGGAGGAAACTCTACCTGATTGATCTTGGCTGGGTTCAGAGGACAC 845
 DB 784 CCATTTCCCGAGCGAGGAGGAAACTCTACCTGATTGATCTTGGCTGGGTTCAGAGGACAC 843
 QY 846 CGGCGCAGGCAACAAGGGGCTTCGGCTAAAAAGAGGTGGAGCCATCAACACCTCCCTG 905
 DB 844 CGGCGCAGGCAACAAGGGGCTTCGGCTAAAAAGAGGTGGAGCCATCAACACCTCCCTG 903
 QY 906 TTTGCTCTGGGCAAGTGGTAGTGGCTGAATCAGGGGCTCCCTCGTGTACCTTATCGG 965
 DB 904 TTTGCTCTGGGCAAGTGGTAGTGGCTGAATCAGGGGCTCCCTCGTGTACCTTATCGG 963

Qy 966 GACAGCAAGCTCACTCGCCTATTTCAGAGACTCTCTGGTGGCTCAGCCACAGTATCCTT 1025
 Db 964 GACAGCAAGCTCACTCGCCTATTTCAGAGACTCTCTGGTGGCTCAGCCACAGTATCCTT 1023
 Qy 1026 ATTGCCCAACATTGCCCTTGAGAGAGCGTCTTACCTAGACACAGTCTCCGCACTCAACTTT 1085
 Db 1024 ATTGCCCAACATTGCCCTTGAGAGAGCGTCTTACCTAGACACAGTCTCCGCACTCAACTTT 1083
 Qy 1086 GCTGCCAGGTCACAGGAGGTGATCAATTG 1114
 Db 1084 GCTGCCAGGTCACAGGAGGTGATCAATCG 1112

RESULT 4
 BT007888
 LOCUS
 DEFINITION Synthetic construct Homo sapiens kinesin-like 4 mRNA, partial cds.
 ACCESSION BT007888
 VERSION BT007888.1 GI:30584614
 KEYWORDS FLI CDNA.
 SOURCE
 ORGANISM
 HUMAN
 REFERENCE
 AUTHORS
 TITLE
 Cloning of human full-length CDSs in BD Creator(TM) System Donor

vector
 1 (bases 1 to 1998)
 2 (bases 1 to 1998)
 Kainline,N., Chen,X., Rolf,A., Halleck,A., Hines,L., Eisenstein,S.,
 Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y.,
 Phelan,M. and Farmer,A.
 Direct Submission
 Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow
 Circle, Palo Alto, CA 94303, USA
 This CDS clone is a part of a collection of human full length
 expression clones generated by BD Biosciences Clontech and the
 Harvard Institute of Proteomics. Each CDS has been cloned in two
 forms: with and without stop-codon (to allow fusion with C-terminal
 tag). The CDS has been directionally cloned using BD In-Fusion(TM)
 cloning system between the SalI and HindIII sites of the pDNR-DUAL
 vector. Additional sequences in the clone: 'ACC' after SalI site
 and before 'Arg' to provide kozak consensus sequence; 'GG' after
 last codon and before HindIII site to maintain reading frame.
 Clone distribution: <http://bioinfo.clontech.com/orfclones>.

FEATURES
 source
 1..1998
 /organism="synthetic construct"
 /mol_type="mRNA"
 /db_xref="taxon:32630"
 /clone="GH00245L1.0"
 /clone_lib="BD Creator (TM) CDS Library derived from MGC
 collection"
 /lab_host="DH5alpha Tl resistant"
 /note="Vector: pDNR-Dual"
 1..>1998
 /note="Mutations: 1997:Stop->Leu"
 /codon_start=1
 /transl_table=11
 /product="Homo sapiens kinesin-like 4"
 /protein_id="P36560.1"
 /db_xref="GI:30584615"

CDS
 1
 translation="MAAGGTQRRREMAAASAAISAGRCRLSKIGATRRPPARV
 RVAVLRFPVDTGASPPCVGRMGDSCSLFIANNRHQETLKYOFDAFYGERSTOOD
 IYAGVQPIRLHGGQNASVAYGPTGAGTKHTMLGSPQPGVPIPRALMDLLQLTRE
 EGAERPWALSYWLYIYQVKVLDLIDPASGLVIREDCRGNILIPGLSQKPISSF
 ADDEHFLPASNRNVTGATRLNQRSSRSHVLLVKVQDRLAPPROREGKLYLIDLA
 GSDNRNTPGKRLKESGAINTSFLVKVVDALNOGLPVPYRDSKLTLLQDLSLG
 GSANITLAPERRFYDITVSNLFAARKEVINRPTNESHQPHALGPVKLSQKE
 LLGPEAKRANGPEEEETGSPPEPMAAPASQKLSPLQKSSMDPAMLERLLSLDRL
 ASQSGQAPLLSTPKRERMVLMKTVEEKDLIERLTKQKLEAKMLAQKABEKNHC

PTMLRPLSHRTVTGAKPLKAVVMPLOLIOEAAASPNABIHILKNKRKRKLESLDAL
 EPEBAEDCWELQISPELLAHGROKILDLINSGSARDLSRLORIGPKKQALIVGREL
 HGFQSVEDLERVEGITGQMESFLKANILGLAAGORCGASL"
 Query Match 99.3%; Score 1107.4; DB 12; Length 1998;
 Best Local Similarity 99.9%; Pred. No. 2.6e-276;
 Matches 1108; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 6 GCCGCGGGCGGCTCGACGACGAGAGGCGACGCGAGATGGCGCAGCTTCAGCGGGCGG 65
 Db 4 GCCGCGGGCGGCTCGACGACGAGAGGCGACGCGAGATGGCGCAGCTTCAGCGGGCGG 63
 Qy 66 ATCTCAGAGAGTGGTGGCTGCTGCTGCTAAGCAAGATTGGAGCTACTCGTCTGCCTCC 125
 Db 64 ATCTCAGAGAGTGGTGGCTGCTGCTGCTAAGCAAGATTGGAGCTACTCGTCTGCCTCC 123
 Qy 126 GCTCGCGTAAAGGTGGCTGTCGACCTGCGGCGCAATTTGTGGATGGAACAGCGGAGCAAGT 185
 Db 124 GCTCGGTAAAGGTGGCTGTCGACCTGCGGCGCAATTTGTGGATGGAACAGCGGAGCAAGT 183
 Qy 186 GATCCCCCTGTGTGCGGGCATATGACAGTGTCTCTAGAGATTGCTTAATCTGGAGAAC 245
 Db 184 GATCCCCCTGTGTGCGGGCATATGACAGTGTCTCTAGAGATTGCTTAATCTGGAGAAC 243
 Qy 246 CACCAGGAGACTCTCAAAATACCAAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAG 305
 Db 244 CACCAGGAGACTCTCAAAATACCAAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAG 303
 Qy 306 GACATCTATGACAGTTTCAGTGCAGCCCATCTTAAGGCACCTTGTGAAAGGCGAGAATGCC 365
 Db 304 GACATCTATGACAGTTTCAGTGCAGCCCATCTTAAGGCACCTTGTGAAAGGCGAGAATGCC 363
 Qy 366 AGTGTGCTTCCCTATGAGCCACAGAGCTGGGAAGCGCACACAATGCTGGCGAGCCCA 425
 Db 364 AGTGTGCTTCCCTATGAGCCACAGAGCTGGGAAGCGCACACAATGCTGGCGAGCCCA 423
 Qy 426 GAGCAACCTGGGCTGATCCCGCGGCTCTCATGAGACTCTCTGAGAGCTCAACAAGGAGGAG 485
 Db 424 GAGCAACCTGGGCTGATCCCGCGGCTCTCATGAGACTCTCTGAGAGCTCAACAAGGAGGAG 483
 Qy 486 GGTGCCGAGGCGCGGCATATGGGCGCTTCTGTGCATCATGCTTCTAGAGATCTACCAG 545
 Db 484 GGTGCCGAGGCGCGGCATATGGGCGCTTCTGTGCATCATGCTTCTAGAGATCTACCAG 543
 Qy 546 GAGAGGATATTAGACCTCTGAGACCTGCTTCGSGAGACCTGTAATCCGAGAGACTGC 605
 Db 544 GAGAGGATATTAGACCTCTGAGACCTGCTTCGSGAGACCTGTAATCCGAGAGACTGC 603
 Qy 606 CGGGGGAATATCTGATTCCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTT 665
 Db 604 CGGGGGAATATCTGATTCCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTT 663
 Qy 666 GAGCGGCATCTCTCTGCGCAGCGAGTGGAAATCGAGCTGTAGAGCCACCGGCTCAACAG 725
 Db 664 GAGCGGCATCTCTCTGCGCAGCGAGTGGAAATCGAGCTGTAGAGCCACCGGCTCAACAG 723
 Qy 726 CGCTCTCCCGCAGTCTGCTGCTGCTCAAGGTGACACAGCGGGAAACGTTTGGCC 785
 Db 724 CGCTCTCCCGCAGTCTGCTGCTGCTCAAGGTGACACAGCGGGAAACGTTTGGCC 783
 Qy 786 CCAATTCGCCAGCGAGAGGAAACCTCTA CTTGATTGACTTGGCTGGGTGAGAGCAAC 845
 Db 784 CCAATTCGCCAGCGAGAGGAAACCTCTA CTTGATTGACTTGGCTGGGTGAGAGCAAC 843
 Qy 846 CGGGCGCAGGCAACAGGCGCTTCGCTAAAGAGAGTGGAGCCATCAACCTCCCTG 905
 Db 844 CGGGCGCAGGCAACAGGCGCTTCGCTAAAGAGAGTGGAGCCATCAACCTCCCTG 903
 Qy 906 TTTGCTCTGGCAAGTGGTAGATGCTGCTGAATCAGGGCCTCCCTCGTGTACCTTATCGG 965
 Db 904 TTTGCTCTGGCAAGTGGTAGATGCTGCTGAATCAGGGCCTCCCTCGTGTACCTTATCGG 963

```
QY 966 GACAGCAAGCTACTCGCCTATTGACGAGCTCTCTGGGTGGCTCAGCCACAGTATCCCTT 1025
|
|
|
Db 964 GACAGCAAGCTACTCGCCTATTGACGAGCTCTCTGGGTGGCTCAGCCACAGTATCCCTT 1023
|
|
|
QY 1026 ATTGCCAACATTGCCCTCAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT 1085
|
|
|
Db 1024 ATTGCCAACATTGCCCTCAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT 1083
|
|
|
QY 1086 GTGCCAGGTCCAAAGGAGGTGATCAATTG 1114
|
|
|
Db 1084 GTGCCAGGTCCAAAGGAGGTGATCAATCG 1112
|
|
|
RESULT 5
AR304057 2097 bp DNA linear PAT 12-JUN-2003
LOCUS Sequence 34 from patent US 6544766.
ACCESSION AR304057
VERSION AR304057.1 GI:31692955
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS 1 (bases 1 to 2097)
TITLE Beraud,C., Ohashi,C., Sakowicz,R., Vaisberg,E., Wood,K. and Yu,M.
JOURNAL Human kinesins and methods of producing and purifying human
FEATURES Patent: US 6544766-A 34 08-APR-2003;
Location/Qualifiers
source 1..2097
/molecule="unknown"
/organism="genomic DNA"
ORIGIN
Query Match 99.3%; Score 1107.4; DB 6; Length 2097;
Best Local Similarity 99.9%; Pred. No. 2.6e-276;
Matches 1108; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 6 GCCCGGGCGGCTCGACGACGAGCGGACGCGAGATGGCGGAGCTTCAGCGGGCGG 65
|
|
|
Db 28 GCCCGGGCGGCTCGACGACGAGCGGACGCGAGATGGCGGAGCTTCAGCGGGCGG 87
|
|
|
QY 66 ATCTCAGAGCTGTGCTGTGCTGCTAAGCAAGATTGGAGCTACTCGTCTCACTCCA 125
|
|
|
Db 88 ATCTCAGAGCTGTGCTGTGCTGCTAAGCAAGATTGGAGCTACTCGTCTCACTCCA 147
|
|
|
QY 126 GCTCCGCTAAGGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 185
|
|
|
Db 148 GCTCCGCTAAGGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 207
|
|
|
QY 186 GATCCCTCTGTGCTGGGGCATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAC 245
|
|
|
Db 208 GATCCCTCTGTGCTGGGGCATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAC 267
|
|
|
QY 246 CACCAAGAGACTCTCAATACCAGTTTGTATGCTTCTATGGGAGAGAGTACTCAGCAG 305
|
|
|
Db 268 CACCAAGAGACTCTCAATACCAGTTTGTATGCTTCTATGGGAGAGAGTACTCAGCAG 327
|
|
|
QY 306 GACATCTATGAGGTTTCACTGAGCGGCTCTTCAAGGCTGCTGAGAGGAGGAGTCC 365
|
|
|
Db 328 GACATCTATGAGGTTTCACTGAGCGGCTCTTCAAGGCTGCTGAGAGGAGGAGTCC 387
|
|
|
QY 366 AGTGTGCTTGGCTATGGACCCACAGAGCTGGGAGAGCGACACAAATGCTGGGAGCCCA 425
|
|
|
Db 388 AGTGTGCTTGGCTATGGACCCACAGAGCTGGGAGAGCGACACAAATGCTGGGAGCCCA 447
|
|
|
QY 426 GAGCAACTGGGGTGAATCCCGGGCTCTCAATGAGCTCTCTGAGCTTCAAGGAGGAG 485
|
|
|
Db 448 GAGCAACTGGGGTGAATCCCGGGCTCTCAATGAGCTCTCTGAGCTTCAAGGAGGAG 507
|
|
|
QY 486 GGTCCGAGGGCGGCGGCTTCTTGTCTACCATGCTTCTACCTAGAGATCTACCAAG 545
|
|
|
Db 508 GGTCCGAGGGCGGCGGCTTCTTGTCTACCATGCTTCTACCTAGAGATCTACCAAG 567
|
|
|
```

```
QY 546 GAGAAGGTATTAGACCTCTCGACCTCTGCTTCCGGAGACCTGCTAATCCGAGAGACTGC 605
|
|
|
Db 568 GAGAAGGTATTAGACCTCTCGACCTCTGCTTCCGGAGACCTGCTAATCCGAGAGACTGC 627
|
|
|
QY 606 CGGGGGAATATCTCTGATTCGGGTCTCTCCAGAGACCCATCAGTAGCTTTGCTGATTTT 665
|
|
|
Db 628 CGGGGGAATATCTCTGATTCGGGTCTCTCCAGAGACCCATCAGTAGCTTTGCTGATTTT 687
|
|
|
QY 666 GAGCGGCACTTCTCGCAGCCAGTCGAATCGACTGTAGGAGCCACCCGGCTCAACCAG 725
|
|
|
Db 688 GAGCGGCACTTCTCGCAGCCAGTCGAATCGACTGTAGGAGCCACCCGGCTCAACCAG 747
|
|
|
QY 726 CGCTCTCTCCGCACTCATGCTGCTCTCTGCTCAAGGTGGAACAGCGGGAACGTTTGGCC 785
|
|
|
Db 748 CGCTCTCTCCGCACTCATGCTGCTCTCTGCTCAAGGTGGAACAGCGGGAACGTTTGGCC 807
|
|
|
QY 786 CCATTTCCGCGAGGAGAGGAAAACTCTACCTGATGACTTGGCTGGGTGAGAGGACAAC 845
|
|
|
Db 808 CCATTTCCGCGAGGAGAGGAAAACTCTACCTGATGACTTGGCTGGGTGAGAGGACAAC 867
|
|
|
QY 846 CGGCGCACAGCAACAAGGGCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTG 905
|
|
|
Db 868 CGGCGCACAGCAACAAGGGCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTG 927
|
|
|
QY 906 TTTGTCTCGGCAAAAGTGTAGTGCCTGCTGAATCAGGGCTCCCTCGTGTACTTATCGG 965
|
|
|
Db 928 TTTGTCTCGGCAAAAGTGTAGTGCCTGCTGAATCAGGGCTCCCTCGTGTACTTATCGG 987
|
|
|
QY 966 GACAGCAAGCTCACTCGCTATTGTCAGGACTCTCTGGGTGGCTCAGCCACAGTATCCCTT 1025
|
|
|
Db 988 GACAGCAAGCTCACTCGCTATTGTCAGGACTCTCTGGGTGGCTCAGCCACAGTATCCCTT 1047
|
|
|
QY 1026 ATTGCCAACATTGGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT 1085
|
|
|
Db 1048 ATTGCCAACATTGGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT 1107
|
|
|
QY 1086 GCTGCCAGGTCCAAAGGAGTGTCAATTG 1114
|
|
|
Db 1108 GCTGCCAGGTCCAAAGGAGTGTCAANTCG 1136
|
|
|
RESULT 6
AB017430 2097 bp mRNA linear PRI 06-MAR-1999
LOCUS Homo sapiens mRNA for kinesin-like DNA binding protein, complete cds.
ACCESSION AB017430 D38751
VERSION AB017430.2 GI:4519442
KEYWORDS Kid; kinesin-like DNA binding protein; kinesin family.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS 1 (sites)
TITLE Tokai,N., Fujimoto-Nishiyama,A., Toyoshima,Y., Yonemura,S., Teukita,S., Inoue,J. and Yamamoto,T.
JOURNAL Kid, a novel kinesin-like DNA binding protein, is localized to chromosomes and the mitotic spindle
MEDLINE EMO J. 15 (3), 457-467 (1996)
PUBMED 96174806
REFERENCE 8599929
AUTHORS 2 (bases 1 to 2097)
TITLE Tokai-Nishizumi,N. and Edamasu,M.
JOURNAL Submitted (03-SEP-1998) Noriko Tokai-Nishizumi, The University of Tokyo, The Institute of Medical Science, Department of Oncology; 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan (E-mail:tokai@hgc.ims.u-tokyo.ac.jp, Tel:03-5449-5302, Fax:03-5449-5413)
COMMENT D38751:Submitted (02-Nov-1994)
Sequence updated (25-Feb-1999).
FEATURES Location/Qualifiers
```

source	1. .2097	/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="pBskid"	
		/cell_line="MDA-MB453"	
		/cell_type="breast cancer"	
gene	1. .2097	/genes="kid"	
CDS	25. .2022	/gene="Kid"	
		/codon_start=1	
		/product="kinesin-like DNA binding protein"	
		/protein_id="BA33019.2"	
		/db_xref="GI:4519443"	
		/translations="MAAGSTQQRREMAAASAAISGACRCLSKIGATRRPPPARV RVAVRLPFDVGTAGSDPPCVRGMDSCSLEIANRNHQETLKYQDFYGERSTOOD IYAGSVOPILRHLLGQNAVLAAGTGAQKTHMLGSPGPGVIFPALMDLQLTRF EGAEHPWALSVMYSVLEIYOEKVLDDLDPASGDLVIREDCRNILIPGLSOKPISSF ADFERHPLASNRVTGATRLNORSRSHAVLLVKVDORERLAPFRQREKLYLIDLA GSDNRRTGNKRLKESGAINSTSLFVLGKVDALNGDLPRVPRDSKLRLLQDSLG GSAHSILIANIAPERFYLDTSALNPAARSKVINRPFNLSLQPHALGPVKSQKE LLGPPKARARPEEIEIGSPMAAPASQSLPLQKLSMDPAMLELLSLDRLL ASQSGAPLLSTPKERMLMTVEKLEIEIRLKTOKELKAKLAKAEKEKNC PTWLRPLSHRTVGAKLKVAVPLQIOEQASPNAEIHLKNNKRRKRLSLDAL EPEKAEDEWELQISPELLAHGRQKILDLNLSGARDLRSLSQRIKPKKAQLIVGWREL HGPFQVEDDLURVEGITGQWESFLKANIILGLAAGQRCGAS"	
ORIGIN			
Query Match	99.3%;	Score 1107.4;	DB 9; Length 2097;
Best Local Similarity	99.9%;	Pred. No. 2.6e-276;	
Matches 1108;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
Qy	6	GC CGCGGCGGCGCTCGACGCGAGCAGAGCGGACGCGAGATGCGGCGAGCTTCAGCGGCGGCG	65
Db	28	GC CGCGGCGGCGCTCGACGCGAGCAGAGCGGACGCGAGATGCGGCGAGCTTCAGCGGCGGCG	87
Qy	66	ATCTCAGAGCTGCTCGCTGCGCTAAGCAAGATTGGAGCTACTCGTCTCCACTTCCA	125
Db	88	ATCTCAGAGCTGCTCGCTGCGCTAAGCAAGATTGGAGCTACTCGTCTCCACTTCCA	147
Qy	126	GCTCGCTAAGGCTGCTGTCGACTGCGGCAATTTGTGGATGGAACAGCGGAGCAAGT	185
Db	148	GCTCGCTAAGGCTGCTGTCGACTGCGGCAATTTGTGGATGGAACAGCGGAGCAAGT	207
Qy	186	GATCCCCCTGTGTCGGGGATGGACAGCTGCTCTTAGAGATTGCTAACTGAGGAAC	245
Db	208	GATCCCCCTGTGTCGGGGATGGACAGCTGCTCTTAGAGATTGCTAACTGAGGAAC	267
Qy	246	CACGAGGAGCTCTCAATACCACTTTGATGCGCTTCTATGGGAGAGGAGTACTCAGCAG	305
Db	268	CACGAGGAGCTCTCAATACCACTTTGATGCGCTTCTATGGGAGAGGAGTACTCAGCAG	327
Qy	306	GACATCTATGAGGTTTCAGTGCAGCCCATCTTAAGGCACTTGTCTGGAAGGCGAGATGCC	365
Db	328	GACATCTATGAGGTTTCAGTGCAGCCCATCTTAAGGCACTTGTCTGGAAGGCGAGATGCC	387
Qy	366	AGTGTGCTTGCTATGACCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGGCCA	425
Db	388	AGTGTGCTTGCTATGACCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGGCCA	447
Qy	426	GAGCAACTGGGGTGATCCGCGGGCTCTCATGACCTCTCTGAGCTCAACAAGGAGGAG	485
Db	448	GAGCAACTGGGGTGATCCGCGGGCTCTCATGACCTCTCTGAGCTCAACAAGGAGGAG	507
Qy	486	GGTCCGAGGCGCGCCATGGGCGCTTTCTGTCTACCAATGCTTACCTAGAGATCTACGAG	545
Db	508	GGTCCGAGGCGCGCCATGGGCGCTTTCTGTCTACCAATGCTTACCTAGAGATCTACGAG	567
Qy	546	GAGAGGATTAGACCTCTGGAACCTGCTTCGGGAGACCTGGTAATCCGAGAGACTGC	605
Db	568	GAGAGGATTAGACCTCTGGAACCTGCTTCGGGAGACCTGGTAATCCGAGAGACTGC	627
Qy	606	CGGCGGAATATCTCTGATTCCGGGTCTCTCCAGAACCCATAGTAGCTTTGCTGATTTT	665
Db	628	CGGCGGAATATCTCTGATTCCGGGTCTCTCCAGAACCCATAGTAGCTTTGCTGATTTT	687
Qy	666	GAGCGGCACTTCTCTGCCAGCAGTGGAAATCGGACTGTGTAGGAGCCACCCGGCTCAACAG	725
Db	688	GAGCGGCACTTCTCTGCCAGCAGTGGAAATCGGACTGTGTAGGAGCCACCCGGCTCAACAG	747
Qy	726	CGCTCTCTCCGCGAGTCTGCTGTGCTTCTCTGCTCAAGGTGAGCAGCGGAGACGTTTGGCC	785
Db	748	CGCTCTCTCCGCGAGTCTGCTGTGCTTCTCTGCTCAAGGTGAGCAGCGGAGACGTTTGGCC	807
Qy	786	CCATTTTCGCCAGCGAGGAGAAACTCTACCTGATTGACTTTGGCTGGGTGAGAGGACAAAC	845
Db	808	CCATTTTCGCCAGCGAGGAGAAACTCTACCTGATTGACTTTGGCTGGGTGAGAGGACAAAC	867
Qy	846	CGGCGCACAGGCAACAAAGGCGCTTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTG	905
Db	868	CGGCGCACAGGCAACAAAGGCGCTTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTG	927
Qy	906	TTTGTCTCTGGGCHAAAGTGTAGTGGCTGAATCAGGGCTTCCCTGCTGATCCTTATCGG	965
Db	928	TTTGTCTCTGGGCHAAAGTGTAGTGGCTGAATCAGGGCTTCCCTGCTGATCCTTATCGG	987
Qy	966	GACAGCAAGCTCACTCGCTTATTGACAGGACTCTCTGGGTGGGTGAGCCACAGTATCCTT	1025
Db	988	GACAGCAAGCTCACTCGCTTATTGACAGGACTCTCTGGGTGGGTGAGCCACAGTATCCTT	1047
Qy	1026	ATTGCCAACAATTGCGGCTGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT	1085
Db	1048	ATTGCCAACAATTGCGGCTGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT	1107
Qy	1086	GCTGCCAGGTCGAGAGGATGATCAATTG	1114
Db	1108	GCTGCCAGGTCGAGAGGATGATCAATCG	1136
RESULT 7			
BC028155		2117 bp	mRNA linear PRI 29-JUN-2004
LOCUS			
DEFINITION			Homo sapiens kinesin family member 22, mRNA (cdna clone MGC:40049
			IMAGE:5241557), complete cds.
ACCESSION			BC028155
VERSION			BC028155.1
KEYWORDS			GI:20380446
SOURCE			MGC
ORGANISM			Homo sapiens (human)
			Homo sapiens
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
			Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE			1 (bases 1 to 2117)
AUTHORS			Strausberg R.L., Feingold, E.A., Grouse, L.H., Dege, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zebberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.C., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Pahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
TITLE			Generation and initial analysis of more than 15,000 full-length
JOURNAL			human and mouse cDNA sequences
PUBMED			Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
REFERENCE			12477932
			2 (bases 1 to 2117)

Db	1089	GCTCCAGGTCACGAGGTCATCATCG	1117		
RESULT 8	BC004352	2134 bp	linear	PRI 29-JUN-2004	
LOCUS	BC004352				
DEFINITION	Homo sapiens kinesin family member 22, mRNA (cdna clone MGC:1573 IMAGE:3535435), complete cds.				
ACCESSION	BC004352				
VERSION	BC004352.1	GI:13279307			
KEYWORDS	MGC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 2134)				
	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusik, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Teichgraber, S., Carninci, P., Frange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Pahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16999-16903 (2002)				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 2134)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: DCTD/DFP cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Anara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhur, Parvaneh Saedi, JR Santos, Angelique Schnerch, Ursula Skalska, Duane Smalhus, Jeff Stott, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Rob Holt, Marco Marra.				
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL Plate: 8 Row: F Column: 1 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6453817.				
FEATURES	Location/Qualifiers				
source	1..2134 /organism="Homo sapiens" /mol_type="mRNA"				
	gene	/db_xref="taxon:9606" /clone="MGC:1573 IMAGE:3535435" /tissue_type="Lung, small cell carcinoma" /clone_lib="NIH MGC 7" /lab_host="DH10B-R" /note="Vector: pOTB7" 1..2134 /gene="KIF22" /note="synonyms: KID, OBP, OBP-1, OBP-2, KNSL4" /db_xref="LocusID:3835" /db_xref="MIM:603213" 20..2017 /gene="KIF22" /codon_start=1 /product="kinesin family member 22" /protein_id="AAH04352.1" /db_xref="GI:13279308" /db_xref="LocusID:3835" /db_xref="MIM:603213" /translation="NRAGSTOORREMAAASAAISGAGRCRLSKI GATRRPPARV RVAVLRPFVDTAGASPPCVGRGMDSCSLEIANRNHOETLKYOFDAFVGRSTOOD IVAGSVQPIRLHLLSQNASVLAAYGTGAGKTHTMLGSPQGVPIPRALMDLLQITRE EAGAEKPMALSVTMSYLEIQEKVLDDLPASGDLVIREDCRNLILIPGLSKYLISSF ADFERHFLPSKRNRTVGATRLNQRSSRAVLLVKVDQRLAPRQREGKYLIDSLA GSFEDNRTGNKRLKESGAINTSFLVGLKVVDALNQGLPRVYPRQSKLITRLQDSLG GSAHSLTIANIAPERRFYLDVTSALNFAARSKEVINRPTNESLOPHALGPVKLSQKE LIGPEAKARGPBESEEGSPMAAPASOKLSPLOKLSMDPAMERLLSLDLRL ASGSGGAPLLSTPKRRNVLMKTVVEKDLERLKTQKELEAKMLAQKASEKENHC PTMLRPLSHRTVGAKPLKAVVPLQLIQOASPNASIEIHLKNGKRKKLESIDLAL EPEKAEDCWEIQI9PELHAGRKILDLINEGSARDLRLSLQRI9GPKKAQLIVGREL HGPFQVEDLERVEGTGQMESFLKANILGLAAGQRCGAS"			
	CDS	Query Match 99.3%; Score 1107.4; DB 9; Length 2134; Best Local Similarity 99.9%; Pred. No. 2.6e-276; Matches 1108; Conservative 0; Mismatches 1; Indels 0; Gaps 0; QY 6 GCCCGCGCGCGCTCGACGCGAGAGCGGCGACGCGAGATGGCGGCGACTTCAGCGCGCGG 65 DB 23 GCCCGCGCGCGCTCGACGCGAGAGCGGCGACGCGAGATGGCGGCGCGGCG 82 QY 66 ATCTCAGGAGCTGGTGGTGGCTTAAGCAAGATTGGAGCTACTCGTCTCACTTCCA 125 DB 83 ATCTCAGGAGCTGGTGGTGGCTTAAGCAAGATTGGAGCTACTCGTCTCACTTCCA 142 QY 126 GCTCGCGTAAAGGCTGGCTGGGACCTGGGCGGCTTTGGATGGAACAGCGGAGCAAGT 185 DB 143 GCTCGCGTAAAGGCTGGCTGGGACCTGGGCGGCTTTGGATGGAACAGCGGAGCAAGT 202 QY 186 GATCCCCCTGTGTGGCGGCGATGGACAGCTGCTCTTAGAGATTGCTAACTGAGGAAC 245 DB 203 GATCCCCCTGTGTGGCGGCGATGGACAGCTGCTCTTAGAGATTGCTAACTGAGGAAC 262 QY 246 CACGAGAGACTCTCAAAATACCAAGTTGATGCTTCTTAGGGGAGAGGAGTACTCAGCAG 305 DB 263 CACGAGAGACTCTCAAAATACCAAGTTGATGCTTCTTAGGGGAGAGGAGTACTCAGCAG 322 QY 306 GACATCTATGCGAGTTTCAGTGCGAGCCATCTCTAAGGCACCTTGTGGAAGGCGAGATGCC 365 DB 323 GACATCTATGCGAGTTTCAGTGCGAGCCATCTCTAAGGCACCTTGTGGAAGGCGAGATGCC 382 QY 366 AGTGTGCTTGCCTATGGACCCACAGGAGCTGGGAAGACGACACAATGTGGGAGGCCCA 425 DB 383 AGTGTGCTTGCCTATGGACCCACAGGAGCTGGGAAGACGACACAATGTGGGAGGCCCA 442 QY 426 GAGCAACTGGGGTGATGCCCGGGGCTCTCTAGGACTCTCTGAGCTCAAGGGAGGAG 485 DB 443 GAGCAACTGGGGTGATGCCCGGGGCTCTCTAGGACTCTCTGAGCTCAAGGGAGGAG 502 QY 486 GGTGCCGAGGCGCGCCATGGGCGCTTTCTGTGCACCATGTCTTACCTAGAGATCTACAG 545 DB 503 GGTGCCGAGGCGCGCCATGGGCGCTTTCTGTGCACCATGTCTTACCTAGAGATCTACAG 562			

QY 546 GAGAGGTATTAGACCTCTGGACCTCTGGACCTCTGGAGAGACCTGTAAATCCGAGAGACTGC 605
 Db 563 GAGAGGTATTAGACCTCTGGACCTCTGGACCTCTGGAGAGACCTGTAAATCCGAGAGACTGC 622
 QY 606 CGGGGGAATATCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTT 665
 Db 623 CGGGGGAATATCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTT 682
 QY 666 GAGCGGCACTTCTGCGACGACGTCGAAATCGGACTGTAGGAGCCACCGGGCTCAACCCAG 725
 Db 683 GAGCGGCACTTCTGCGACGACGTCGAAATCGGACTGTAGGAGCCACCGGGCTCAACCCAG 742
 QY 726 CGCTCTCCCGAGTCATGCTGTGCTCTGCTGCTCAAGGTGGAACAGCGGAAAGCTTTGGCC 785
 Db 743 CGCTCTCCCGAGTCATGCTGTGCTCTGCTGCTCAAGGTGGAACAGCGGAAAGCTTTGGCC 802
 QY 786 CCATTTTCGCGAGGAGGGAACCTCTACCTGATTTGACTTTGGCTGGGTTCAGAGGACAC 845
 Db 803 CCATTTTCGCGAGGAGGGAACCTCTACCTGATTTGACTTTGGCTGGGTTCAGAGGACAC 862
 QY 846 CGCGGCACAGGCAACAGGGCCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTG 905
 Db 863 CGCGGCACAGGCAACAGGGCCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTG 922
 QY 906 TTTGCTCTGGGCAAAAGTGTAGATGCGTGAATCAGGGCCCTCCCTGTGTACTTATCGG 965
 Db 923 TTTGCTCTGGGCAAAAGTGTAGATGCGTGAATCAGGGCCCTCCCTGTGTACTTATCGG 982
 QY 966 GACAGCAAGCTCACTCGCTATTGAGGACTCTCTGGGTGGCTCAGCCACAGTAGTTCCTT 1025
 Db 983 GACAGCAAGCTCACTCGCTATTGAGGACTCTCTGGGTGGCTCAGCCACAGTAGTTCCTT 1042
 QY 1026 ATTGCCAATGTCCTCGAGAGAGCTTCTACCTAGACACAGTCTCCGACCTCAACTTT 1085
 Db 1043 ATTGCCAATGTCCTCGAGAGAGCTTCTACCTAGACACAGTCTCCGACCTCAACTTT 1102
 QY 1086 GCTGCCAGTCCAGGAGGTGATCAATTG 1114
 Db 1103 GCTGCCAGTCCAGGAGGTGATCAATCG 1131

RESULT 9
 AR210054
 LOCUS AR210054 1041 bp DNA linear PAT 20-JUN-2002
 DEFINITION Sequence 5 from patent US 6387644.
 ACCESSION AR210054
 VERSION AR210054.1 GI:21512185
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1041)
 AUTHORS Beraud,C.
 TITLE Motor proteins and methods for their use
 JOURNAL Patent: US 6387644-A 5 14-MAY-2002;
 FEATURES
 source 1..1041
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Query Match 93.1%; Score 1038; DB 6; Length 1041;
 Best Local Similarity 100.0%; Pred. No. 2.6e-258;
 Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 78 GGTCTGTCGGCTAAGCAAGATTGGAGTCTCTGTCGTCACCTCCAGCTCCGCTAAGG 137
 Db 4 GGTCTGTCGGCTAAGCAAGATTGGAGTCTCTGTCGTCACCTCCAGCTCCGCTAAGG 63
 QY 138 GTGGCTGTGCGACTGGGCAATTGTGGATGAAACAGCGGGAGCAAGTAGTCCCGCTGT 197
 Db 64 GTGGCTGTGCGACTGGGCAATTGTGGATGAAACAGCGGGAGCAAGTAGTCCCGCTGT 123

QY 198 GTGCGGGCATGAGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAACCAACAGGAGACT 257
 Db 124 GTGCGGGCATGAGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAACCAACAGGAGACT 183
 QY 258 CTCAAATACAGTTTGTATGCTCTCTATGCGGAGAGAGTACTCAGCAGGACATCTATGCA 317
 Db 184 CTCAAATACAGTTTGTATGCTCTCTATGCGGAGAGAGTACTCAGCAGGACATCTATGCA 243
 QY 318 GGTTCAGTGCACGCCCATCTTAAGGCACTTGTCTGGAAGGCGCAATGCCAGTGTGCTTGGC 377
 Db 244 GGTTCAGTGCACGCCCATCTTAAGGCACTTGTCTGGAAGGCGCAATGCCAGTGTGCTTGGC 303
 QY 378 TATGAGCCACAGGAGCTGGGAGAGCGACACAATGCTGGGAGCCCGAGAGAACTCTGGG 437
 Db 304 TATGAGCCACAGGAGCTGGGAGAGCGACACAATGCTGGGAGCCCGAGAGAACTCTGGG 363
 QY 438 GTGATCCCGCGGGCTCTCATGACCTCTGCGAGCTCAAAAGGAGAGGGTGCAGAGGC 497
 Db 364 GTGATCCCGCGGGCTCTCATGACCTCTGCGAGCTCAAAAGGAGAGGGTGCAGAGGC 423
 QY 498 CGGCCATGGGCCCTTCTGTCAACCATGCTTACCTAGAGATCTACCAGGAGAGGTATTA 557
 Db 424 CGGCCATGGGCCCTTCTGTCAACCATGCTTACCTAGAGATCTACCAGGAGAGGTATTA 483
 QY 558 GACCTCTGGACCCCTGCTTCGGGAGACCTGTGTAATCCGAGAGAGACTGCCGGGGAATATC 617
 Db 484 GACCTCTGGACCCCTGCTTCGGGAGACCTGTGTAATCCGAGAGAGACTGCCGGGGAATATC 543
 QY 618 CTGATTCGGGCTCTTCCGAGAGCCCATCAGTAGCTTTGCTGATTTTGAAGCGGCACTTC 677
 Db 544 CTGATTCGGGCTCTTCCGAGAGCCCATCAGTAGCTTTGCTGATTTTGAAGCGGCACTTC 603
 QY 678 CTGCGAGCCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACGAGGCTCTCTCCGC 737
 Db 604 CTGCGAGCCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACGAGGCTCTCTCCGC 663
 QY 738 AGTCATGCTGCTCTCTGCTCAAGGTGGACAGCGGGAACGTTTGGCCCCATTTGCGCCAG 797
 Db 664 AGTCATGCTGCTCTCTGCTCAAGGTGGACAGCGGGAACGTTTGGCCCCATTTGCGCCAG 723
 QY 798 CGAGAGGAAAACTCTACTGATTCATTTGGCTGGGTCAAGGACCAACCGGCGCACAGGC 857
 Db 724 CGAGAGGAAAACTCTACTGATTCATTTGGCTGGGTCAAGGACCAACCGGCGCACAGGC 783
 QY 858 AACAGGGCCCTTCGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTGCTCTGGGC 917
 Db 784 AACAGGGCCCTTCGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTGCTCTGGGC 843
 QY 918 AAAGTGTAGATGCGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGAGACAGCAAGCTC 977
 Db 844 AAAGTGTAGATGCGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGAGACAGCAAGCTC 903
 QY 978 ACTGCTCTATTGCAAGGACTCTCTGGGTGGCTCAGCCCAAGTATCTTATTTGCCAACATT 1037
 Db 904 ACTGCTCTATTGCAAGGACTCTCTGGGTGGCTCAGCCCAAGTATCTTATTTGCCAACATT 963
 QY 1038 GCGGCTGAGAGACGCTTCTACTAGACAGTCTCGGACCTCAACTTTGCTGCGCAGGTC 1097
 Db 964 GCGGCTGAGAGACGCTTCTACTAGACAGTCTCGGACCTCAACTTTGCTGCGCAGGTC 1023
 QY 1098 AAGGAGGTGATCAATTGA 1115
 Db 1024 AAGGAGGTGATCAATTGA 1041

RESULT 10
 AR210055
 LOCUS AR210055 1464 bp DNA linear PAT 20-JUN-2002
 DEFINITION Sequence 7 from patent US 6387644.
 ACCESSION AR210055
 VERSION AR210055.1 GI:21512186
 KEYWORDS
 SOURCE Unknown.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 5 Row: m Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21704181.

FEATURES

Source

source	1. .2086 /organism="Mus musculus" /mol_type="mRNA" /strain="FVB/N" /db_xref="taxon:10090" /clone="MGC:6456 IMAGE:2615715" /issue_type="Mammary tumor, C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse." /clone_11b="NCI CGAP Mam6" /lab_host="DH10B" /note="Vector: pCwV-SPORT6"
gene	1. .2086 /gene="Kif22" /note="synonym: Kid" /db_xref="LocustID:110033" /db_xref="MGI:109233" 20. .2002 /gene="Kif22" /codon_start=1 /product="kinesin family member 22" /protein_id="AAH03427.1" /db_xref="GI:13097360" /db_xref="LocustID:110033" /db_xref="MGI:109233" /translation="MSLRATCSQQRREMASTSGRCVSKGGLGRPPPLARVRAVR LRPPMDETEAKELPCVRAIGTSCGLEVANWKKYQETLKTKYQFDAPYGEKSTQQQVYVGS VQPLRLHLLGSQCNVLAIVDPTGAGKTHTMLGSPQEPQVIFPRMADGKTLQARRESAEG RPMDSVAMSGLYLDLDDPASGLDVIREDRCGNILI PGLTKQGITPSDFEQ HFPASNRAGVATRLNQRSRSRSHVLLVKVDQERLTPFRQREGKLYLIDLAGSDN RRTYQNGILRKESGAINTSFLVKGVVADLNQGLPRI PYRDSKLTLLDSDSLGGSASH ILLANI APERRFYODTISALNFTASRKEVINRPTNESLQPHALAPVKLSQELLGGS EAKKAGPESSTGSPSTAPASQKLSLLKLSNMDPAMLENLLSWERLLGSGS QGTPLNTPKRRWNLVVEENKLETRUMKQKELEAKVLAQEAADPREKENTPTFI LQPPASVGTGVAKPLKAVVMYLPQRI QKQRESSNQILKKGPKRKLSPSEAEVK DEDYWEVQISPELLHGRKLLDLNLAASOARELSRQICQKKAQILVGRHLHGPPS EVEDLEOVESIGKGVESFLKANLLSLAASOHSPPS"
CDS	

ORIGIN

Query Match	69.9%	Score	779.8	DB	10	Length	2086
Best Local Similarity	83.6%	Pred. No.	2.7e-191				
Matches	880	Conservative	0	Mismatches	173	Indels	0
0;							
QY	59	GGCGGCGATCTCAGGAGCTGGTCTGCTGGCTTAAGCAAGATTGGAGCTACTGTCGTCC	118				
Db	61	GGCGTGGCCACCTCAGGACCTGGCCGCTGTGTAAAGCAAGAGGAGCTTGGCGGCGTCC	120				
QY	119	ACTTCAGCTCGCGTAAAGGTGGCTGTGCATCGCGGCCATTGTGGATGGAAACAGCGGG	178				
Db	121	GCCTCTAGCTCGAGTAAGGGTGGCTGTGCGGCTGCGCCGCTTTATGATGGAGACAGACA	180				
QY	179	AGCAAGTGATCCCCCTGTGTGGGGGCATGGACAGCTGCTCTAGAGATTGCTAACTG	238				
Db	181	AGCGAAGGAGTCCCTGTGTCCGAGCCCATAGACAGCTGCTCTCTTGAATGGCTAACTG	240				
QY	239	GAGCAACACCAAGAGACTCTCAAAATACCAAGTTTGATGCTCTATGGGAGAGGAGTAC	298				
Db	241	GAGAAATACCGAGAGACGCTCAAAATATCATAGTTTGATGCCCTCTATGGCGAAGAGCAC	300				
QY	299	TCAGCAGGACATCTATGTCAGGTTTCAGTGCAGGCCATCCTAAGGCACCTGCTGGAAGGCCA	358				
Db	301	TCAGCAGGAGTCTATGTAGGTTTCAGTACAGCCCATCTTAAGGCATCTGCTGGAAGGCCA	360				
QY	359	GAATGCCAGTGTGCTTTGCCCTATGGACCCACAGGAGCTGGGAAGACGCACACAATGCTGGG	418				
Db	361	GAATGCCAGTGTACTTTGCCCTATGGGCTACTGGGGCAGGGAGACACACAAATGCTGGG	420				
QY	419	CAGCCACAGCAACCTGGGGGTGATCCCGGGGCTCTCATGGACCTCTCTCAGCTCAACAAG	478				

DECIIT 12.

RESOL 12
AC101919/C

ACTUARY

FOCUS DEFINITION

NOT YET JUD

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE	DATE	TIME	LOCATION	STATUS
1. [Illegible]	10/10/2023	14:30	Room 101	Completed
2. [Illegible]	10/11/2023	10:00	Room 102	In Progress
3. [Illegible]	10/12/2023	09:00	Room 103	Not Started
4. [Illegible]	10/13/2023	11:00	Room 104	Completed
5. [Illegible]	10/14/2023	13:00	Room 105	In Progress
6. [Illegible]	10/15/2023	15:00	Room 106	Not Started
7. [Illegible]	10/16/2023	16:00	Room 107	Completed
8. [Illegible]	10/17/2023	17:00	Room 108	In Progress
9. [Illegible]	10/18/2023	18:00	Room 109	Not Started
10. [Illegible]	10/19/2023	19:00	Room 110	Completed

JOURNAL

REFERENCE

AUTHORS

Db	421	CAGCCAGAGCAACCTGGAGTGATTCCTCGGCTCTCATGGACCTCTGCAACTAGCAAG	480
Qy	479	GGAGGAGGTCGCGAGGCGCGCATGGGCCCTTTCTGTCAACCATCTCTTACCTAGAGAT	538
Db	481	GGAGAGAGTGCTGAGGAGCGGCATGGGACGTTTCTGTGCTATGTCCTATTATGAGAT	540
Qy	539	CTACCGAGAGAGGTATTAGACCTCTCGGACCTGCTCGGAGACCTGCTAATCCGAGA	598
Db	541	CTACCGAGAGAGGTATTAGACCTCTTGATCTGATCAGGAGACTCGTGATCCGCA	600
Qy	599	AGACTGCGGGGGAATATCTGATTCGGGTCTCTCCAGAAAGCCCATAGTAGCTTTGC	658
Db	601	AGACTGCGGGGGAATATCTGATTCGAGGCTTCACAGAAAGCCCATCCAGCTTCTC	660
Qy	659	TGATTTTGAGCGCACTTCTGCGCAGCGAGTCAAAATCGAATGATAGGAGCCACCGGCT	718
Db	661	TGACTTCGAGCAGCACTTCTTCAGCAGTAGTAAATCGAGCTGTAGGAGCCACCGGCT	720
Qy	719	CAACACGCGCTCTCCGCGAGTCATCTGTGCTCTCGGTCAAGGTGGACCGCGGAACG	778
Db	721	TAACCGAGCTCTTCCGCTAGTCAACGAGTGTCTTGGTCNAGGTAGATCAGGCTGAACG	780
Qy	779	TTTGCGCCCATTTGCGCAGGAGAGGAGAACTCTACTGATTCGCTGGCTGGGTGAG	838
Db	781	TTTGACTCTCATTTTGGCAGCGGAGAGAACTCTACTTATGATTGGCTGGTTTCTG	840
Qy	839	GGACACCGGCGCACAGGCAACAGAGGCTCTCGGCTAAAGAGAGTGAGGCCATCAAC	898
Db	841	GGACACCGGCGCACAGGCAACAGAGGCTCTCGGCTAAAGAGAGTGAGGCCATCAAC	900
Qy	899	CTCCCTGTTTGTCTCGGCGAAAGTGTAGATGCGCTGAATCAGGGCTCTCCCTGCTGACC	958
Db	901	CTCCCTGTTTGTCTCGGCGAAAGTGTAGATGCGCTGAATCAGGGCTCTCCCTGCTGACC	960
Qy	959	TTATCGGGACAGCAAGCTCACTGCGCTATTCAGGAGCTCTCTGGGTGGCTCAGCCACAG	1018
Db	961	ATACCGGGACAGCAAGCTCACTGCTGCTGCGAGGCTCTCTGGGAGGCTCAGCTCATAG	1020
Qy	1019	TATCTTATTGCCAATTTGCCCTGAGAGAGCTCTTACCTAGACAGAGCTCTCGGCT	1078
Db	1021	CATCTCTATTTGCCAATTTGCCCTGAGAGAGCTCTTACCTAGACAGAGCTCTCGGCT	1080
Qy	1079	CAACTTTGCTGCGCAGGTCAGAGGAGGTGATCAA	1111
Db	1081	GAATTTGCTGCGCAGGTCAGAGGAGGTGATCAA	1113
RESULT 12			
AC101919/c			
LOCUS			
DEFINITION	AC101919	151041 bp	DNA linear HTG 14-MAR-2003
ACCESSION	Mus musculus clone RP24-158A6	WORKING DRAFT SEQUENCE	4 unordered pieces.
VERSION	AC101919		
KEYWORDS	AC101919.4	GI:28951338	
SOURCE	HTG; HTGS PHASE1; HTGS DRAFT.		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Mus musculus		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
JOURNAL	1 (bases 1 to 151041)		
TITLE	Birren, B., Nusbaum, C. and Lander, E.		
REFERENCE	Mus musculus, clone RP24-158A6		
AUTHORS	2 (bases 1 to 151041)		
	Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choquel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A., Cooke, P., DeArnell, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hago, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,		

Db 65130 TCTGACTTCGAGAACGACTTCCTTCTAGTCAGTAGAAATCGAGTCTTAGGAGCTACTCTG 65071
QY 717 CTCAACAGCGGCTCCCGCAGTCATGCTGCTCTCTGTCAGGTGACGACCGGGAA 776
Db 65070 CTTAACAGCGGCTCTGCGGTAGTCATGACGCGCTCTTGTCAAGGTAGATCAGCATGAA 65011
QY 777 CGTTTGGCCCAATTCGCCAGCGAGAGGAGAACTCTACTGATTCATTCGCTGGCTGGTCA 836
Db 65010 CGTTTGACTCCATTCCTGCGCAGTGGGAAGGAAACTCTACCTATTGCTTGGCTGGTCA 64951
QY 837 GAGGACAACC-----GCGCACAGGCAACAGGCGCTTCGGCTAAAGAGAGTGGAGCCAT 892
Db 64950 GAGGACAACCGTCAGTCGACAGGCAACAGGCGCTTCGCTCAAGAGAGTGGAGCCAT 64891
QY 893 CCAACACTCCCTGTTGCTCTGCGCAAGTGGTAGATGCGCTGAATCAGGCGCTCCCTCG 952
Db 64890 CCAACACTCCCTGTTGCTCTGCGCAAGTGGTAGATGCGCTGAATCAGGCGCTCCCTCG 64831
QY 953 TGTACCTTATCGGAGCAGCAAGCTCACTCGCCTATTGCGAGGACTCTCTGGTGGCTCAGC 1012
Db 64830 TATACCATACGAGCAGCAGGCTCACTCGTCTGTTGCAAGACTCTCTGAGAGGCTCAGC 64771
QY 1013 CCACAGTATCCTTATTCGCAACTTCGCCCTGAGACGCTTCTACCTAGACACAGTCTTC 1072
Db 64770 TCATAGCATCCTCATTTGTCAGCATTCCTCTGAAAGATGTTTTTACCTGGATACATCTC 64711
QY 1073 CGCACTCAACTTGTGCTGCGCAGGTCGCAAGGAGTGATCAA 1111
Db 64710 AGCATTAACCTTCACTGCTAGGTCGCAAGGAGTGATTA 64672

RESULT 13

AC101752/c
LOCUS AC101752
DEFINITION Mus musculus chromosome 1, clone RP24-344C18, complete sequence.
ACCESSION AC101752.14 GI:46518680
VERSION AC101752.14
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 164759)
REFERENCE
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE Mus musculus chromosome 1, clone RP24-344C18
JOURNAL Unpublished
2 (bases 1 to 164759)
REFERENCE
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cook, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K.,
Lamares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,
Olivier, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Riebeck, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS

3 (bases 1 to 164759)
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collymore, A., Cook, A., Cooke, P., Corum, B., DeArrellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neil, D., Olivier, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL

Submitted (17-MAR-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 164759)

REFERENCE
AUTHORS

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, S., Arachchi, H.M., Barna, N., Bastien, V.,
Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J.,
Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B.,
DeArrellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D.,
Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
O'Connor, T., O'Donnell, P., O'Neil, D., Olivier, J., Peterson, K.,
Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
Zimmer, A. and Zody, M.

TITLE
JOURNAL

Submitted (23-APR-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 23, 2004 this sequence version replaced gi:45504330.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

COMMENT

Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@broad.mit.edu
----- Project Information
Center project name: L17344
Center clone name: 344_C_18

FEATURES
source

Location/Qualifiers
1..164759
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/map="1"
/clone="RP24-344C18"
/clone_lib="RPMI-24 Male Mouse BAC"
repeat_region 1..1655
/rpt_family="L1_MM"

```

1. .4
/note="clone boundary
clone_end:SF6
site:MboI"
1767. .1787
/rpt family="AT_rich"
2115. .2157
/rpt family="AT_rich"
2236. .2258
/rpt family="AT_rich"
3252. .3420
/rpt family="Lx7"
3464. .3588
/rpt family="Lx7"
3605. .3666
/rpt family="CT-rich"
3673. .3731
/rpt family="TTC)n"
3732. .3911
/rpt family="TTC)n"
4000. .4180
/rpt family="CT-rich"
4202. .4363
/rpt family="Lx7"
4529. .4765
/rpt family="MTE"
7234. .7442
/rpt family="B3"
8445. .8509
/rpt family="T-rich"
complement (8512. .8591)
/rpt family="MLT11"
9336. .9632
/rpt family="MLT13"
complement (9800. .11396)
/rpt family="L1_MM"
complement (11397. .17527)
/rpt family="L1F"
complement (18235. .18777)
/rpt family="Lx"
18775. .18890
/rpt family="Lx"
19260. .19949
/rpt family="RMER19A"
20118. .20297
/rpt family="MLT1D"
complement (20386. .20639)
/rpt family="Lx2"
20634. .20839
/rpt family="Lx4"
20876. .21793
/rpt family="RLTRI13D"
21799. .21907
/rpt family="RMER13A"
21918. .22433
/rpt family="Lx4"
22591. .22651
/rpt family="TG)n"
22733. .24486
/rpt family="L1_MM"
24494. .24850
/rpt family="L1_MM"
27139. .27166
/rpt family="AT_rich"
27737. .27834
/rpt family="MLT1H"
28328. .28506
/rpt family="MLT1K"
28513. .28541
/rpt family="AT_rich"
28813. .28858
/rpt family="AT_rich"
29627. .30186

```

```

repeat_region
/rpt family="L1_MM"
30465. .30578
/rpt family="T-rich"
complement (30581. .34457)
/rpt family="L1_MM"
complement (34455. .35895)
/rpt family="L1"
complement (36134. .36736)
/rpt family="Lx4"
36749. .36920
/rpt family="Lx2"
37617. .37751
/rpt family="(TA)n"
37958. .37986
/rpt family="(GAA)n"

Query Match 61.1%; Score 681.4; DB 10; Length 164759;
Best Local Similarity 79.7%; Pred. No. 8.9e-166;
Matches 844; Conservative 0; Mismatches 206; Indels 9; Gaps 3;

Qy 57 GCGCGCGGATCTCAGAGCTGCTCGCTGCGGCTAAGCAAGATGGAGCTACTCGTCGT 116
Db 146137 GCSTCAGCCACCTCAGGACCGGGCTTTTGT---ATAAGCAAGGTAGGATTTGGCTGCGT 146081
Qy 117 CCACCTCCAGCTCGCGTAAGSGTGGCTGCGACTGCGGCCCATTTGTGGATGGAACAGCG 176
Db 146080 CTGCTTCTAGCTGGAGTAAGGTGGCTATATATTGTGCTGTTTGGATGAAGAGACA 146021
Qy 177 GGAGCAAGTGATCCCCCTGTGTGCGGGGATGACAGCTGCTCTCTAGAGATTGCTAAC 236
Db 146020 GAAGCGAAGGAGCACCCCTGTGTACGAGCCATACAGAGCTGCTCTCTGAAGTGCCTAAC 145961
Qy 237 TGGAGGAACCAACAGGAGACTCTCAAAATACAGTTTGCCTTCTATGGGGAGAGAGT 296
Db 145960 TGGAGGAATAACAGGAGACTATCAAAATATCAGTTTGCCTTCTATGGAGAGAGAGC 145901
Qy 297 ACTCAGCAGGACATCTATGAGGTTTCTGAGTGCAGGCCATCTTAAGGCACTTCTCGAAGGG 356
Db 145900 ACTCAGTGGACATCTATGAGTTTGTAGTACAGCCCATCTAAGAACTTGTCTGAAGGG 145841
Qy 357 CAGAAATCCAGTGTGCTTGTGCTATGAGCCCAACAGGAGCTGGGAGAGACACAAATGCTG 416
Db 145840 CAGAAATGCCAGTGTACTTGTCTATGAGCCGACTGGGCGAGGGAAGA--CACACAATGTTG 145783
Qy 417 GGCAGCCCAACAGCAGTGGGGTGTATCCCGGGGCTCTCATGGACCTCTCTGACCTCACA 476
Db 145782 GGCAGCCCAACAGCAGTGGAGTGTCTCTCAGACTCTCCAGACCTCTCTGACCTAACA 145723
Qy 477 AGGAGGAGGCTGCGGAGGCGCGCCCATGGGCCCTTTCTGTACCATGTCTTACCTAGAG 536
Db 145722 AGGAGAGAGATGCTGAGGCGTGGCCATGGGACATTTCTGTGTTTGTGCTATTTAGAT 145663
Qy 537 ATCTACAGGAGAGGATTTAGACCTCTCTGAGCCCTCTCTCGGAGACCTTGGTAAATCCGA 596
Db 145662 CTCTAACAGGAAATAGTATTAGACCTCTTGGATCTGCAATCAGAAAGACCTAGTGTCCGA 145603
Qy 597 GAAGACTGCGGGGGAATATCTTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTT 656
Db 145602 GAAGGCTGACGAGGGAATCTTGATTCAGGCGCTCA CAGAAAGCCCATCAGTAGTTTC 145543
Qy 657 GCTGATTTTGAAGGCACTTCTGCGCAGCCAGTCCGAATCGGACTGTAGGAGCCACCCGG 716
Db 145542 TCTGACTTCGAGAGCACTTCTTCTAGTCAGTAGAATCGAGTTGTAGGAGCTACTCTG 145483
Qy 717 CTCAACAGCGCTCTCCCGAGTCTGCTGTGCTCTCTGCTCAAGGTGAGCCAGCGGAA 776
Db 145482 CTTAACAGCGCTCTTTCGCTAGTCTGCGGCTCTTGTGTCAGGCTAGTCAAGGTAGATCAGATCAA 145423
Qy 777 GGTTCGCCCCATTTCCGCGAGGAGGGAATCTTACCTGATTTGACTTGGCTGGGTCA 836
Db 145422 CGTTTGATCTCCATTTCCCGAGGGAATCTTACCTTATTTGATTTGGCTGGTTCA 145363
Qy 837 GAGGACAACC---GGCGCACAGGCAACAAAGGGCTTTGGCTAAAGAGAGTGGAGCCAT 892

```

```

Db 145362 GAGGACAAACGTCAGTCGACAGCAACGAGGCATTCCTCAAGAGAGTGGAGCCAT 145303
QY 893 CAACACCTCCCTGTTTCTGCGCAAGTGTAGATGCGTGAATCAGGGCTCCCTCG 952
Db 145302 CAACACCTCCCTGTTTCTGCGCAAGTGTAGATGCGTGAATCAGGGCTCCCTCA 145243
QY 953 TGTACCTTATCGGACAGCAAGCTCACTCGCCCTATTGACAGGACTCTCTGGTGGTTCAGC 1012
Db 145242 TATACCATACGAGCAGCAGGCTCACTCGTCTGTTGCAAGACTCTCTGAGAGGCTCAGC 145183
QY 1013 CCACAGATCCCTATTGCAACATGCGCCCTGAGACAGCCTTCTACCTAGACAGATCTC 1072
Db 145182 TCATAGCATCCTCATGTCAGCATTCCTCTGAAGATGTTTTTACCTGGATACATCTC 145123
QY 1073 CGCACTCAACTTGTGTCGCGAGGTCGAAGGAGGTGATCAA 1111
Db 145122 AGCATTAACCTTCACTGCTAGGTCGAAGGAGGTGATTA 145084

```

```

RESULT 14
AC101908/c
LOCUS
DEFINITION
AC101908
AC101908.4 GI:44199137
HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
MUS musculus
MUS musculus (house mouse)
REFERENCE
AUTHORS
TITLE
MUS musculus chromosome 1, clone RP24-275J1
Unpublished
2 (bases 1 to 297639)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
Lamazares,R., Landers,T., Lehotzky,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Straelus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 297639)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafes,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,

```

```

Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (18-JUN-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 27, 2004 this sequence version replaced gi:31880232.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
----- Project Information
Center project name: Li7732
Center clone name: 275_J_1
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

```

```

1 63747: contig of 63747 bp in length
* 63748 63847: gap of 100 bp
* 63848 73892: contig of 10045 bp in length
* 73893 73992: gap of 100 bp
* 73993 154450: contig of 80458 bp in length
* 154451 154550: gap of 100 bp
* 154551 167808: contig of 13258 bp in length
* 167809 167908: gap of 100 bp
* 167909 176711: contig of 8803 bp in length
* 176712 176811: gap of 100 bp
* 176812 233500: contig of 56889 bp in length
* 233501 233600: gap of 100 bp
* 233601 241222: contig of 7622 bp in length
* 241223 241323: gap of 100 bp
* 241323 297639: contig of 56317 bp in length.
Location/Qualifiers
1. 297639
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/map="1"
/clone="RP24-275J1"
/clone_lib="RPCI-24 Male Mouse BAC"

```

FEATURES

source

ORIGIN

```

Query Match 58.7%; Score 654.2; DB 2; Length 297639;
Best Local Similarity 78.1%; Pred. No. 1e-158;
Matches 827; Conservative 0; Mismatches 223; Indels 9; Gaps 3;
QY 57 GCGCGGGCGATCTCAGGAGCTGGTGGCTGGCTGGCTGAAGCAAGATTGGAGCTACTGCTGCT 116
Db 1205 GCGTCAGCCACTCAGGACCGGCTTTGT---ATAAGCAGGTAGGATTGGCTGCTGCT 1149
QY 117 CCACCTCAGCTCGGTAAGGGTGGCTGTGGCACTGCGGCCATTTGTGGATGGAACAGCG 176
Db 1148 CTGCTCTTAGCTGGAGTAAGGGTGGCTGTATATAATTTGCTGCTGTTTGTGGATGAAGAGACA 1089

```


Qy	177	GGAGCAAGTGATCCCCCTGTGTGGGGCATGGACAGCTGCTCTAGAGATTGCTAAC	236
Ds	1088	GAAGGAGGAGGACCCCTGTGTAGGACCATAGACAGCTGCTCTTGAAGTGGCTAAC	1029
Qy	237	TGGAGGAACCAACGAGGAGACTCTCAATACAGTTTGTATGCTTATGGGAGAGAGT	296
Ds	1028	TGGAGGAATACGAGGAGACTATCAATATCAGTTTGTATGCTTCCATGGAGAGAGAGC	969
Qy	297	ACTCAGCAGGACATCTATGAGGTTCTAGTGGAGCCCATCTTAAGCACTTCTGGAGGG	356
Ds	968	ACTCAGTGGGACATCTATGAGGTTTAGTACAGCCCATCTTAAGCACTTCTGGAGGG	909
Qy	357	CAGAAATGCACTGCTTCCCTATGAGCCACAGGAGCTGGGAGAGACACAAATGCTG	416
Ds	908	CAGAAATGCACTGCTTCCCTATGAGCCACAGGAGCTGGGAGAGAGACACAAATGCTG	851
Qy	417	GGCAGCCAGAGCAACCTGGGGTATCCCGGGCTCTCATGGACCTCTGACACTGCTTAC	476
Ds	850	GGCAGCCAGAGCAACCTGGGGTATCCCGGGCTCTCATGGACCTCTGACACTGCTTAC	791
Qy	477	AGGAGGAGGGTCCGAGGGCCGAGGCTTCTGTCACCATGCTTACCTAGAG	536
Ds	790	AGGAGGAGGGTCCGAGGGCCGAGGCTTCTGTCACCATGCTTACCTAGAG	731
Qy	537	ATCTACAGGAGAGGATATTAGACTCTGAGCCCTGCTTCCGAGAGCTTGGTAATCCGA	596
Ds	730	CTCTAACAGGAAATAGTATTAGACTCTTGTATCTTCTATAGTGTACCTAATGTCGA	671
Qy	597	GAAGACTGCCGGGGAATATCTGATCCGGGTCTTCCAGAGAGCCCATCAGTAGCTTT	656
Ds	670	CGGCGAGGAGGAGGAAATCTGATCCAGGCTTCCAGAGAGCCCATCAGTAGCTTT	611
Qy	657	GCTGATTTGAGCGGCACTTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	716
Ds	610	TCTGACTTCGAGAGGAGCACTTCTTCTAGTCACTAGGAGGAGGAGGAGGAGGAGG	551
Qy	717	CTCAACAGGCGCTCTCCCGAGTCATGCTGTGCTCTGCTGCTGCTGCTGCTGCTGCTG	776
Ds	550	CTTAACAGGCGCTCTTGGCGTAGTCATGCGGCTCTTGGTCAAGGATAGATCAGCATGA	491
Qy	777	CGTTTGGCCCATTTGCGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	836
Ds	490	CGTTTGGCTCATTTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	431
Qy	837	GAGCAGACCC---GGGCGCAGGCAACAGGCGCTTGGCTAAAGAGAGTGGAGCAT	892
Ds	430	GAGCAGACCCCTGAGTGGCAGGCAACAGGCGCTTGGCTAAAGAGAGTGGAGCAT	371
Qy	893	CAACACCTCTCTGTTTGTCTGGGCAAGTGGATGAGTGGCTGAATCAGGCGCTTCTCTG	952
Ds	370	CAACACCTCTCTGTTTGTCTGGGCAAGTGGATGAGTGGCTGAATCAGGCGCTTCTCTG	311
Qy	953	TGTAACCTTATCGGAGCAGGAGCTCATCTGCGCTTATGAGGAGCTCTCTGGGTGCTCAGC	1012
Ds	310	TATACCATACGAGCAGGAGCTCATCTGCGCTTATGAGGAGCTCTCTGGGTGCTCAGC	251
Qy	1013	CCACAGTATCTTATTTGCCAATTTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1072
Ds	250	TCATAGATCTCTTATTTGCCAATTTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	191
Qy	1073	CGCACTCAACTTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1111
Ds	190	AGCATTAACCTTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	152

RESULT 15
 AC102127
 LOCUS AC102127
 DEFINITION Mus musculus chromosome 1, clone RP23-202A19, complete sequence.
 ACCESSION AC102127
 VERSION AC102127.11
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 196674)
 Birren,B., Nusbaum,C. and Lander,E.
 Mus musculus chromosome 1, clone RP23-202A19
 Unpublished
 2 (bases 1 to 196674)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,
 Brown,A., Camarata,J., Campiano,A., Chang,J., Charaz,B.,
 Choepel,Y., Collings,M., Collins,S., Collymore,A., Cook,A.,
 Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S.,
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
 Lamazares,R., Landers,T., Lehotsky,J., Levine,R., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Melarm,J.,
 Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollard,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Vial,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE

JOURNAL

REFERENCE
AUTHORS

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 196674)

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
 Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
 Bloom,T., Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J.,
 Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
 Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
 Erickson,J., Fato,S., Ferreira,P., FitzGerald,M., Gage,D.,
 Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
 Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
 Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
 Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
 MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
 McCarthy,M., Melarm,J., Meneus,L., Mihova,T., Mlenga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
 Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
 Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Vial,R.,
 Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
 Zimmer,A. and Zody,M.

TITLE

JOURNAL

REFERENCE
AUTHORS

Submitted (10-APR-2004) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 196674)

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
 Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
 Bloom,T., Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J.,
 Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
 Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
 Erickson,J., Fato,S., Ferreira,P., FitzGerald,M., Gage,D.,
 Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
 Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
 Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
 Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
 MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
 McCarthy,M., Melarm,J., Meneus,L., Mihova,T., Nicol,R., Norbu,C.,
 Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,

Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
 Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R.,
 Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
 Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (01-MAY-2004) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 1, 2004 this sequence version replaced gi:46358225.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/MIT Center for Genome Research

Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@broad.mit.edu
 ----- Project Information
 Center project name: L18071
 Center clone name: 202_A_19

FEATURES

source

Location/Qualifiers

1..196674
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="1"
 /map="1"
 /clone="RP23-202A19"
 /clone_lib="RPCI-23 Female Mouse BAC"
 1..4229
 /note="wgs end extension
 clone end:SP6"
 1..193
 /rpt_family="MTA"
 614..746
 /note="single clone coverage"
 701..706
 /note="<30 qual single clone coverage"
 1208..1291
 /rpt_family="(CTA)n"
 1571..1842
 /rpt_family="RLTRETn MM"
 complement(1974..2057)
 /rpt_family="B4"
 3219..3319
 /rpt_family="GA-rich"
 3368..3448
 /rpt_family="(CA)n"
 3626..3688
 /rpt_family="(CAAA)n"
 3925..4374
 /rpt_family="Lx"
 4230..4235
 /note="clone boundary
 clone end:SP6
 site:EcoRI"
 4383..4411
 /rpt_family="AT-rich"
 4445..4496
 /rpt_family="(TCTCTG)n"
 4509..4638
 /rpt_family="(TG)n"
 4741..4784
 /rpt_family="(TG)n"
 4785..4832
 /rpt_family="(TA)n"
 4833..4955
 /rpt_family="(TG)n"
 5398..5431
 /rpt_family="(GAAA)n"

repeat_region complement(5434..5561)
 /rpt_family="RMER15"
 repeat_region complement(6185..6272)
 /rpt_family="RSIN1"
 repeat_region 6516..6570
 /rpt_family="(T)n"
 repeat_region complement(6571..6823)
 /rpt_family="L1_MM"
 repeat_region 6821..6939
 /rpt_family="L1_MM"
 repeat_region 7007..7285
 /rpt_family="Lx5"
 repeat_region 7292..7848
 /rpt_family="Lx5"
 repeat_region 8187..8286
 /rpt_family="BSIN1"
 repeat_region complement(8823..9061)
 /rpt_family="B4A"
 repeat_region complement(9243..9815)
 /rpt_family="L1"
 repeat_region 9832..9861
 /rpt_family="(A)n"
 repeat_region 9862..10368
 /rpt_family="Lx2"
 repeat_region complement(10373..11559)
 /rpt_family="L1_MM"
 repeat_region 11555..12196
 /rpt_family="Lx2"
 repeat_region complement(13221..13381)
 /rpt_family="B3A"
 13437..13578
 /rpt_family="B1_MM"
 repeat_region 15222..15254
 /rpt_family="AT-rich"
 repeat_region complement(15315..15605)
 /rpt_family="RMER20"
 repeat_region complement(16008..16517)
 /rpt_family="RMERS"
 repeat_region 16673..16713
 /rpt_family="(TG)n"
 repeat_region 16809..16866
 /rpt_family="(TG)n"
 repeat_region complement(16922..17052)
 /rpt_family="B4"
 repeat_region 17101..17490
 /rpt_family="MLT2B1"
 repeat_region 17725..17762
 /rpt_family="(GGAA)n"
 Query Match 46.9%; Score 522.8; DB 10; Length 196674;
 Best Local Similarity 74.4%; Pred. No. 1.3e-124;
 Matches 726; Conservative 0; Mismatches 242; Indels 8; Gaps 5;
 QY 143 TGTGGACTGGCGCCATTGTGGATGGAAACAGCGGAGCAAGTGTATCCCTGTGTGG 202
 Db 101632 TTTATGACTATGCCATTTATGATGAAGAAAAAGAAAGGATGCTTGTGTAA 101691
 QY 203 GGGCATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGACCAACAGGAGACTCTCA 262
 Db 101692 AGGCATATAAAGCTGCTCTCTTGGAGTGCCACCTTATGAGAAATCTTATGAGAAATCTAA 101751
 QY 263 ATACAGTTTCATGCTTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGCAGGTTTC 322
 Db 101752 ATATCAGTTTAAATGTCTTTTATTTAGAGAAAGCACTCAGCAGAACATCTAAGTAGGTTTC 101811
 QY 323 AGTGAGCCCATCTTAAGGCACCTTGTCTGGAAGGGCAGAAATCCAGTGTGCTTGCCTATGG 382
 Db 101812 AGTACAGCTATCTGTGGCCTTGTGTGGTGGCAGAAATACCACTGTGCTTACATATGG 101871
 QY 383 ACCCAGGAGCTGGGAAGACGCACAAATCTCTGGGAGCCC-----AGAGCACTGGGG 438
 Db 101872 GACTATTGAGCAGGGAAGACACATATAATACTGGGTAGCCCCATGGGAACAACATGGAG 101931

Qy 439 TGATCCCGGGCTCTCATGACCTCTGAGCTCAACAGGAGGAGGTGCCGAGGCC 498
 Db 101932 TGATTCCTTAGCCCTCTCATAGACCTCTGAGATAACAAGGAGAAAGTCTGAAGGCT 101991
 Qy 499 GGCCATGGGCCCTTCTGTCCACCATGCTTACCTAGAGATCTACCCAGGAGAAGGTATTAG 558
 Db 101992 GGCCATGGGACATTTCTGTAGGATGCTTATTTAGAGGTCTACCCAGGAAAAGGTATTAG 102051
 Qy 559 ACCTCTGGACCTGTCTTCCGGAGACCTGGTAATCCGAGAGACTGCCGGGGGAATATCC 618
 Db 102052 ACCTCTTATATCTGTCATCAGGAGACCTAGTGATATGAGAAGATTGTGAGGGAACATCC 102111
 Qy 619 TGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGAGGGCACTTCC 678
 Db 102112 TGATTCAGTCTCTCACCAGGACCCCAATCTGCTTCTCTGACTTTTGAGCAGTACTT-C 102170
 Qy 679 TGCCAGGCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACCCAGCGCTCTCTCCCGCA 738
 Db 102171 TGCCAGCCTGTAGAAATCAAGTTGTAGGAGTACTCTGCTTAAGCAGTGTCTTACCATA 102230
 Qy 739 GTCATGTGTCTCTCTGGTCAAGGTGACACAGCGGGAACGTTTGGCCCCCATTTGCCCAGC 798
 Db 102231 -TCATTCAATGCTCTTTATCAAGGTGGATCAGCATGGACATTTGACTCTGTTTCACTAGT 102289
 Qy 799 GAGAGGGAAGAACTCTACTGATTGACTTTGGCTGGGTGAGAGCAACCCGGCGCACAGGCA 858
 Db 102290 GAAAGGAAAGAACTCTACCATATTGATTGGTTGGCTCAGAGGACCAACAACA-AGGA 102348
 Qy 859 ACAAGGCGCTTCGGCTAAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCA 918
 Db 102349 ACCAGTGCATTCAACTCAAGAGAGTGGCGCCATCAACACCT-GCTTCTTGTAAGTGGCA 102407
 Qy 919 AAGTGGTAGATGCGCTGAATCAGGCGCTCCCTCGTGTAACCTTATCGGAGACAGCAAGTCA 978
 Db 102408 AAGTGGAGTTGCATCATATTAGGGCTCCCTCTCTATATATCTGGGACAGAAAAATCA 102467
 Qy 979 CTCGCTATTGCGAGACTCTCTGGGTGGCTCAGGCCACAGTATCTTTATTGCCAACATTG 1038
 Db 102468 TTTGCTATTGCCAGACTCTCTGGAAGCTCAGCTCATACATCATCTTACCACACTG 102527
 Qy 1039 CCCCTGAGAGAGCTTTCTACCTAGACACAGTCTCCGCACTCAACTTTGTCGCCAGGTCCA 1098
 Db 102528 CTCCTGAAAGACATTTTCTTGGAAATCATTTGAGCATTTGAACCTTCACTGCTAGTCCA 102587
 Qy 1099 AGGAGGTGATCAATTG 1114
 Db 102588 AGGAGCTGATTAATTG 102603

Search completed: November 10, 2004, 13:17:54
 Job time : 4934.36 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 13:18:18 ; Search time 781.387 Seconds
(without alignments)
10087.639 Million cell updates/sec

Title: US-10-797-893-7

Perfect score: 1464

Sequence: 1 atgggtcgctgtcggttaag.....accattgtcccaatgtga 1464

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3611042 seqs, 2692057975 residues

Total number of hits satisfying chosen parameters: 7222084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1464	100.0	1464	17	US-10-797-893-7
2	1461	99.8	1538	17	US-10-797-893-3
3	1447	98.8	2099	16	US-10-334-143-102
4	1443.2	98.6	2104	9	US-09-925-300-288
5	1038.4	70.9	1041	17	US-10-797-893-5
6	1035.4	70.7	1115	17	US-10-797-893-1
7	480	32.8	491	9	US-09-960-253-131
8	376.4	25.7	386	10	US-09-803-719-1937
9	365	24.9	24525	13	US-10-087-192-1180
10	265	18.1	464	10	US-09-918-995-31705
11	262.4	17.9	24291	13	US-10-087-192-1177
12	233.2	15.9	531	15	US-10-029-386-10603

13	230.4	15.7	232	15	US-10-029-386-24303	Sequence 24303, A
14	185.6	12.7	2095	16	US-10-108-260A-249	Sequence 249, App
15	156.8	10.7	3374	16	US-10-287-226-301	Sequence 301, App
16	138.6	9.5	4108	9	US-09-883-096-1	Sequence 1, Appli
17	137.2	9.4	2675	16	US-10-287-226-305	Sequence 305, App
18	136.6	9.3	1152	9	US-09-883-096-6	Sequence 6, Appli
c 19	134	9.2	2607	17	US-10-437-963-1460	Sequence 1460, App
c 20	127.6	8.7	482	16	US-10-152-319A-498	Sequence 498, App
21	123.2	8.4	2698	16	US-10-425-114-25129	Sequence 25129, A
22	121.8	8.3	3661	16	US-10-311-642-1	Sequence 1, Appli
23	121.6	8.3	3034	18	US-10-425-115-148256	Sequence 148256,
24	121.4	8.3	1014	9	US-09-883-096-3	Sequence 3, Appli
25	113.6	7.8	987	16	US-10-311-642-3	Sequence 313, App
26	112.2	7.7	4292	16	US-10-287-226-313	Sequence 720, App
27	110.2	7.5	740	14	US-10-076-555-720	Sequence 384, App
28	107	7.3	259	14	US-10-076-555-384	Sequence 828, App
29	107	7.3	657	14	US-10-076-555-828	Sequence 56548, A
30	106.2	7.3	1998	17	US-10-437-963-56548	Sequence 135546,
31	101	6.9	2264	18	US-10-425-115-135546	Sequence 586, App
32	101	6.9	3258	16	US-10-302-172-586	Sequence 23, Appli
33	100.2	6.8	3657	16	US-10-336-472-23	Sequence 57, Appli
34	100.2	6.8	3657	16	US-10-236-417-57	Sequence 1087, Ap
35	98.2	6.7	4724	15	US-10-172-118-1087	Sequence 1087, Ap
36	98.2	6.7	4724	16	US-10-342-887-1087	Sequence 201, App
37	94.4	6.4	2217	15	US-10-104-047-201	Sequence 22135, A
38	91.4	6.2	3390	16	US-10-425-114-22135	Sequence 2359, Ap
39	91.4	6.2	3407	18	US-10-739-930-2359	Sequence 46568, A
40	91.4	6.2	3433	18	US-10-425-115-46568	Sequence 34, Appli
41	90.4	6.2	4913	15	US-10-133-937-34	Sequence 34, Appli
42	90.4	6.2	4913	16	US-10-159-563-34	Sequence 27001, A
43	88.8	6.1	1295	15	US-10-369-493-27001	Sequence 762, App
44	88.4	6.0	3159	15	US-10-104-047-762	Sequence 1053, Ap
45	86.2	5.9	6927	15	US-10-172-118-1053	

ALIGNMENTS

RESULT 1
US-10-797-893-7
; Sequence 7, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10797,893
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Human
US-10-797-893-7

Query Match 100.0%; Score 1464; DB 17; Length 1464;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1464; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGGTA 60
Db 1 ATGGGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGGTA 60
Qy 61 AGGGTGTGCTGTGCGCTACTGCGGCCATTGTGGATGGAAACAGGGAGCAAGTATCCCCC 120
Db 61 AGGGTGTGCTGTGCGCTACTGCGGCCATTGTGGATGGAAACAGGGAGCAAGTATCCCCC 120

QY 121 TGTGTGGGGGATGAGACAGTCTCTCTAGAGATTGCTAACTGGAGGAACACAGGAG 180
 Db 121 TGTGTGGGGGATGAGACAGTCTCTCTAGAGATTGCTAACTGGAGGAACACAGGAG 180
 QY 181 ACTCTCAATACACAGTTTCATGCTCTTATGGGAGAGAGTACTCAGCAGACATCTAT 240
 Db 181 ACTCTCAATACACAGTTTCATGCTCTTATGGGAGAGAGTACTCAGCAGACATCTAT 240
 QY 241 GCAGGTTTCAGTGCAGGCCCATCTTAAAGGCACTTGTGGAAGGCGAGATCCAGTGTGCTT 300
 Db 241 GCAGGTTTCAGTGCAGGCCCATCTTAAAGGCACTTGTGGAAGGCGAGATCCAGTGTGCTT 300
 QY 301 GCTTATGACCCACAGAGAGTGGGAAGACGCAACAATGCTGGGCGAGCCGAGCAACT 360
 Db 301 GCTTATGACCCACAGAGAGTGGGAAGACGCAACAATGCTGGGCGAGCCGAGCAACT 360
 QY 361 GGGGTGATCCCGGGGCTCTCATGGACCTCTTACAGCTTCAAGGAGGAGGTCGCGAG 420
 Db 361 GGGGTGATCCCGGGGCTCTCATGGACCTCTTACAGCTTCAAGGAGGAGGTCGCGAG 420
 QY 421 GCGCGGCCATGGGCCCTTCTGTACCATGTCTTACCTAGAGATCTACAGGAGAAGTA 480
 Db 421 GCGCGGCCATGGGCCCTTCTGTACCATGTCTTACCTAGAGATCTACAGGAGAAGTA 480
 QY 481 TTAGACTCTCTGACCTCTTCCGAGACCTGGTAATCGGAGAGACTGCCGGGGAAT 540
 Db 481 TTAGACTCTCTGACCTCTTCCGAGACCTGGTAATCGGAGAGACTGCCGGGGAAT 540
 QY 541 ATCTGATTCGGGGTCTCTCCAGAGGCCATCAGTAGCTTGTCTGATTTGAGCGGCAC 600
 Db 541 ATCTGATTCGGGGTCTCTCCAGAGGCCATCAGTAGCTTGTCTGATTTGAGCGGCAC 600
 QY 601 TTCTGCGACCCAGTCCGAATCGAGTGTAGGAGCCACCGGCTCAACAGGCTCTCTCC 660
 Db 601 TTCTGCGACCCAGTCCGAATCGAGTGTAGGAGCCACCGGCTCAACAGGCTCTCTCC 660
 QY 661 CGCAGTCATGCTCTGCTCTGCTCAAGTGGACGAGGGGAGTTTGGCCCATTTCCG 720
 Db 661 CGCAGTCATGCTCTGCTCTGCTCAAGTGGACGAGGGGAGTTTGGCCCATTTCCG 720
 QY 721 CAGCAGAGGGAATCTTACCTGATTTGACTTGGCTGGGTGAGGACCAACCGGCGACA 780
 Db 721 CAGCAGAGGGAATCTTACCTGATTTGACTTGGCTGGGTGAGGACCAACCGGCGACA 780
 QY 781 GGCAACAAGGGCTCTGGCTTAAAGAGAGTGGAGCCATCAACCTCTGTTTGTCTG 840
 Db 781 GGCAACAAGGGCTCTGGCTTAAAGAGAGTGGAGCCATCAACCTCTGTTTGTCTG 840
 QY 841 GGCAGAGTGTAGATGCTGATCAAGGCTCTGAGGCTCTGCTGATCTTATCGGACAGCAAG 900
 Db 841 GGCAGAGTGTAGATGCTGATCAAGGCTCTGAGGCTCTGCTGATCTTATCGGACAGCAAG 900
 QY 901 CTCACCTGCTTATGACAGACTCTCTGGGTGGCTCAGCCCAAGTATCTTATGCAAC 960
 Db 901 CTCACCTGCTTATGACAGACTCTCTGGGTGGCTCAGCCCAAGTATCTTATGCAAC 960
 QY 961 ATTGCCCCGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTCCAG 1020
 Db 961 ATTGCCCCGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTCCAG 1020
 QY 1021 TCCAGGAGTGTATCAATCGGCTTTTACCAATGAGAGCTGAGCTCATGCTTGGGA 1080
 Db 1021 TCCAGGAGTGTATCAATCGGCTTTTACCAATGAGAGCTGAGCTCATGCTTGGGA 1080
 QY 1081 CCTGTTAAGCTGTCTCAGAAGAATTCCTTGGTCCACAGAGGCAAGAGAGCCGAGGC 1140
 Db 1081 CCTGTTAAGCTGTCTCAGAAGAATTCCTTGGTCCACAGAGGCAAGAGAGCCGAGGC 1140
 QY 1141 CCTGAGGAAGAGGATTTGAGGCCCTGAGCCCATGACGCTCCAGCTCTGCTCCAG 1200
 Db 1141 CCTGAGGAAGAGGATTTGAGGCCCTGAGCCCATGACGCTCCAGCTCTGCTCCAG 1200
 QY 1201 AAACCTCAGCCCCCTTACAGAGCTAAGCAGCATGGACCCGGCCATGCTGGAGCGCTCCTC 1260

Db 1201 AAACCTCAGCCCCCTTACAGAGCTAAGCAGCATGGACCCGGCATGCTGGAGCGCTCCTC 1260
 QY 1261 AGCTTTGGACCGTCTGCTTGCCTCCAGGGGAGCAGGGGGCCCTCTCTTTGAGTACCCCA 1320
 Db 1261 AGCTTTGGACCGTCTGCTTGCCTCCAGGGGAGCAGGGGGCCCTCTCTTTGAGTACCCCA 1320
 QY 1321 AAGCGAGAGCGATGGTGTCTAATGAGACAGTGAAGAGAGAGGACCTAGAGATTGAGAGG 1380
 Db 1321 AAGCGAGAGCGATGGTGTCTAATGAGACAGTGAAGAGAGAGGACCTAGAGATTGAGAGG 1380
 QY 1381 CTTTAAAGCAAGCAAAAAAAGAACTGGAGGCCAAGATGTTGGCCCAAGATGTTGGCCCAAGAGGCTGAGGAAAG 1440
 Db 1381 CTTTAAAGCAAGCAAAAAAAGAACTGGAGGCCAAGATGTTGGCCCAAGATGTTGGCCCAAGAGGCTGAGGAAAG 1440
 QY 1441 GAGAACCAATGTCACCAATGTGA 1464
 Db 1441 GAGAACCAATGTCACCAATGTGA 1464

RESULT 2
 US-10-797-893-3
 ; Sequence 3, Application US/10797893
 ; Publication No. US20040142397A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Berard, Christophe
 ; TITLE OF INVENTION: Novel motor proteins and methods for
 ; FILE REFERENCE: 1044
 ; CURRENT APPLICATION NUMBER: US/10/797,893
 ; CURRENT FILING DATE: 2004-03-09
 ; PRIOR APPLICATION NUMBER: US/09/724,224
 ; PRIOR FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: 09/597,292
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 1538
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-10-797-893-3

Query Match 99.8%; Score 1461; DB 17; Length 1538;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGTGCTGTGCTAAGCAAGATTGGAGCTACTGCTGCTCCACCTCCAGCTCGCGTAAGG 63
 Db 78 GGTGCTGTGCTAAGCAAGATTGGAGCTACTGCTGCTCCACCTCCAGCTCGCGTAAGG 137
 QY 64 GTGGCTGTGCTAAGCAAGATTGGAGCTACTGCTGCTCCACCTCCAGCTCGCGTAAGG 123
 Db 138 GTGGCTGTGCTAAGCAAGATTGGAGCTACTGCTGCTCCACCTCCAGCTCGCGTAAGG 197
 QY 124 GTGGGGGATGAGAGCTGCTCTTAGAGATTGCTAACTGGAGGAACACACAGAGACT 183
 Db 198 GTGGGGGATGAGAGCTGCTCTTAGAGATTGCTAACTGGAGGAACACACAGAGACT 257
 QY 184 CTCAAATACAGTTTGTATGCTTCTATGGGGAGAGAGTACTCAGCAGACATCTATGCA 243
 Db 258 CTCAAATACAGTTTGTATGCTTCTATGGGGAGAGAGTACTCAGCAGACATCTATGCA 317
 QY 244 GGTTCAGTGCAGCCCATCTTAAGGCACTTGTGGAGGGCAGAGTCCAGTGTGCTTGGC 303
 Db 318 GGTTCAGTGCAGCCCATCTTAAGGCACTTGTGGAGGGCAGAGTCCAGTGTGCTTGGC 377
 QY 304 TATGAGCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGCCGAGAGCAACTGGG 363
 Db 378 TATGAGCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGCCGAGAGCAACTGGG 437
 QY 364 GTGATCCCGGGGCTCTCATGACCTCTGAGCTCACAAGGGAGGAGGTCGCGAGGCG 423

Db 438 GTGATCCCGGGCTCTCATGGAAGCTCTGTGAGCTCAAGAGGAGGAGGGTCCGAGGGC 497
Qy 424 CGGCCATGGGCCCTTTCTGTCAACATGCTTTACCTAGAGATCTACCAAGAGAAAGTATTA 483
Db 498 CGGCCATGGGCCCTTTCTGTCAACATGCTTTACCTAGAGATCTACCAAGAGAAAGTATTA 557
Qy 484 GACCTCTGGACCTCTGTTCCGAGAGACCTGTGTAATCCGAGAAGACTGCGGGGGAATATC 543
Db 558 GACCTCTGGACCTCTGTTCCGAGAGACCTGTGTAATCCGAGAAGACTGCGGGGGAATATC 617
Qy 544 CTGATTCGGGCTCTCTCCAGAAGCCATCAGTACCTGCTGATTTGAGCGGCACTTC 603
Db 618 CTGATTCGGGCTCTCTCCAGAAGCCATCAGTACCTGCTGATTTGAGCGGCACTTC 677
Qy 604 CTGCAGCAGTCAAGATCGAGCTGTAGAGCCACCGGCTCAACACAGCGCTCTCTCCGCG 663
Db 678 CTGCAGCAGTCAAGATCGAGCTGTAGAGCCACCGGCTCAACACAGCGCTCTCTCCGCG 737
Qy 664 AGTCATGCTGCTCTCTCCAGAAGCCATCAGTACCTGCTGATTTGAGCGGCACTTC 723
Db 738 AGTCATGCTGCTCTCTCCAGAAGCCATCAGTACCTGCTGATTTGAGCGGCACTTC 797
Qy 724 CGAGAGGAAAGCTCACTACCTGATTTGACTTGGCTGGTCAAGAGCAACCGGCGCACAGC 783
Db 798 CGAGAGGAAAGCTCACTACCTGATTTGACTTGGCTGGTCAAGAGCAACCGGCGCACAGC 857
Qy 784 AACAGGGCCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGC 843
Db 858 AACAGGGCCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGC 917
Qy 844 AAAGTGTAGATGCGCTGATTCAGGGCTCCCTGCTGTACCTTATCGGACAGCAAGCTC 903
Db 918 AAAGTGTAGATGCGCTGATTCAGGGCTCCCTGCTGTACCTTATCGGACAGCAAGCTC 977
Qy 904 ACTGCCCTATTCAGGAGCTCTCGGCTGGCTCAGCCACAGTATTCCTTATGCGCAACAT 963
Db 978 ACTGCCCTATTCAGGAGCTCTCGGCTGGCTCAGCCACAGTATTCCTTATGCGCAACAT 1037
Qy 964 GCCCTCAGAGAGCGCTTCTACCTAGACACAGTCTCGGCTCAGGCTCACTTGTGCGAGTCC 1023
Db 1038 GCCCTCAGAGAGCGCTTCTACCTAGACACAGTCTCGGCTCAGGCTCACTTGTGCGAGTCC 1097
Qy 1024 AAGAGGTGATCAATCGGCTTTTACCAATGAGAGCTCGAGCTCATGCTTGGGACCT 1083
Db 1098 AAGAGGTGATCAATCGGCTTTTACCAATGAGAGCTCGAGCTCATGCTTGGGACCT 1157
Qy 1084 GTTAAGCTGTCTGAGAGAAATGCTTGTGTCACAGAGCAAGAGAGCCGAGGCGCT 1143
Db 1158 GTTAAGCTGTCTGAGAGAAATGCTTGTGTCACAGAGCAAGAGAGCCGAGGCGCT 1217
Qy 1144 GAGGAGAGGAGATTGGAGCCCTGAGCCATGAGCTCCAGCTCTGCTCCAGAAA 1203
Db 1218 GAGGAGAGGAGATTGGAGCCCTGAGCCATGAGCTCCAGCTCTGCTCCAGAAA 1277
Qy 1204 CTCAGCCCTCTACAGAGCTAAGCAGCATGGACCCGCGCTATGCTGGAGCGCTCTCAGC 1263
Db 1278 CTCAGCCCTCTACAGAGCTAAGCAGCATGGACCCGCGCTATGCTGGAGCGCTCTCAGC 1337
Qy 1264 TTGAGCGCTGTGCTCTCCAGGAGGAGCCAGGGGCGCTCTGTTGAGTACCCCAAG 1323
Db 1338 TTGAGCGCTGTGCTCTCCAGGAGGAGCCAGGGGCGCTCTGTTGAGTACCCCAAG 1397
Qy 1324 CGAGAGCGGATGCTTAATGAGACAGTATGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 1383
Db 1398 CGAGAGCGGATGCTTAATGAGACAGTATGAGAGAGGAGGAGGAGGAGGAGGAGGAG 1457
Qy 1384 AAGACGAGCAAAAGAACTGGAGGCAAGATGTTGGCCAGAGGCTGAGGAAAGGAG 1443
Db 1458 AAGACGAGCAAAAGAACTGGAGGCAAGATGTTGGCCAGAGGCTGAGGAAAGGAG 1517
Qy 1444 AACCATTTGCCCAATGTGA 1464
Db 1518 AACCATTTGCCCAATGTGA 1538

RESULT 3

US-10-334-143-102
; Sequence 102, Application US/10334143
; Publication No. US20040009549A1
; GENERAL INFORMATION:
; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD
; FILE REFERENCE: 038602/1543
; CURRENT APPLICATION NUMBER: US/10/334,143
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: 60/343,169
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 102
; LENGTH: 2099
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-334-143-102

Query Match 98.8%; Score 1447; DB 16; Length 2099;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1458; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 4 GGTGCTGTGCTCGGCTAAGCAAGATTGGAGCTACTGCTGCTCCACCTCCAGCTCGGTAAGG 63
Db 101 GGTGCTGTGCTCGGCTAAGCAAGATTGGAGCTACTGCTGCTCCACCTCCAGCTCGGTAAGG 160
Qy 64 GTGGCTGTGCTGCTCGGCTCGGCTATTTGCTGATGGAACAGCGGGAGCAAGTATCCCCCTGT 123
Db 161 GTGGCTGTGCTGCTCGGCTCGGCTATTTGCTGATGGAACAGCGGGAGCAAGTATCCCCCTGT 220
Qy 124 GTGGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGAAACCAACAGAGACT 183
Db 221 GTGGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGAAACCAACAGAGACT 280
Qy 184 CTCAATATACAGTTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 243
Db 281 CTCAATATACAGTTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 340
Qy 244 GGTTCAGTGCAGCCCATCTTAAGGCACTTCTGGAAGGGGAGAGTGCAGTGTGCTTGGC 303
Db 341 GGTTCAGTGCAGCCCATCTTAAGGCACTTCTGGAAGGGGAGAGTGCAGTGTGCTTGGC 400
Qy 304 TATGGACCCACAGAGAGCTGGGAAAGACGACACAATGCTGGGAGCCGAGAGCAACCTGGG 363
Db 401 TATGGACCCACAGAGAGCTGGGAAAGACGACACAATGCTGGGAGCCGAGAGCAACCTGGG 460
Qy 364 GTGATCCCGGGCTCTCATGGACCTCTCTGAGCTCAAGAGGAGGAGGCTGCGAGGGC 423
Db 461 GTGATCCCGGGCTCTCATGGACCTCTCTGAGCTCAAGAGGAGGAGGCTGCGAGGGC 520
Qy 424 CGGCCATGGGCCCTTTCTGTACCATGCTTTACTAGAGATCTACAGAGAGAGGATTA 483
Db 521 CGGCCATGGGCCCTTTCTGTACCATGCTTTACTAGAGATCTACAGAGAGAGGATTA 580
Qy 484 GACCTCTGAGACCTCTCTCGGAGAGCTGTAATCCGAGAAGACTGCGGGGGAATATC 543
Db 581 GACCTCTGAGACCTCTCTCGGAGAGCTGTAATCCGAGAAGACTGCGGGGGAATATC 640
Qy 544 CTGATTCGGGCTCTCTCCAGAGAGCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTC 603
Db 641 CTGATTCGGGCTCTCTCCAGAGAGCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTC 700
Qy 604 CTGCCAGCAGTGCAGAAATCGGATGTAGAGAGCAACCGGCTCAACAGAGCTCTCCCGC 663
Db 701 CTGCCAGCAGTGCAGAAATCGGATGTAGAGAGCAACCGGCTCAACAGAGCTCTCCCGC 760
Qy 664 AGTCATGCTGTGCTCTCTGCTCAAGTGGACAGGGGAAAGCTTTGGCCCCCATTTTCGCGAG 723

Db 761 AGTCATGCTGCTCTCTGCTCAAGGTGGACGAGGGAACGTTTGGCCCCATTTCGCCAG 820
QY 724 CGAGAGGGAAAACCTCTACCTGATGATGAGCTTGGGTGAGAGCAACCGCGGCGACAGGC 783
Db 821 CGAGAGGGAAAACCTCTACCTGATGATGAGCTTGGGTGAGAGCAACCGCGGCGACAGGC 880
QY 784 AACRAGGGCTTCGGCTAAAGAGAGTGGAGCATCAACACCTCCCTGTTGCTCTGGGC 843
Db 881 AACRAAGGGCTTCGGCTAAAGAGAGTGGAGCATCAACACCTCCCTGTTGCTCTGGGC 940
QY 844 AAGTGGTAGATGCGCTGAATCAGGGCCCTCCCTGCTGTACCTTATCGGGACACGAAGTTC 903
Db 941 AAGTGGTAGATGCGCTGAATCAGGGCCCTCCCTGCTGTACCTTATCGGGACACGAAGTTC 1000
QY 904 ACTCGCTATTGAGAGCTCTCTGGGTGGCTCAGGCCACAGTATCTTTATTCGCAACATT 963
Db 1001 ACTCGCTATTGAGAGCTCTCTGGGTGGCTCAGGCCACAGTATCTTTATTCGCAACATT 1060
QY 964 GCCCTGAGAGAGCTTCTACCTAGACAGAGTCTCGGCACTCAACTTTGCTGCGAGGTCC 1023
Db 1061 GCCCTGAGAGAGCTTCTACCTAGACAGAGTCTCGGCACTCAACTTTGCTGCGAGGTCC 1120
QY 1024 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTCGAGCCTCATGCTTGGGACCT 1083
Db 1121 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTCGAGCCTCATGCTTGGGACCT 1180
QY 1084 GTTAAGCTCTCAGAAAAGAAATGCTTGTGTCACAGAGGCAAGAGAGCCCGAGGCCCT 1143
Db 1181 GTTAAGCTCTCAGAAAAGAAATGCTTGTGTCACAGAGGCAAGAGAGCCCGAGGCCCT 1240
QY 1144 GAGGAGAGAGATTTGGGAGCCCTGAGCCATGAGAGCTCGAGCCTCTGCTCCAGAAA 1203
Db 1241 GAGGAGAGAGATTTGGGAGCCCTGAGCCATGAGAGCTCGAGCCTCTGCTCCAGAAA 1300
QY 1204 CTCAGCCCCCTCAGAAAGCTTAAGCAGCATGAGAGCCCGGCCATGCTGGAGCGCTCTCTCAGC 1263
Db 1301 CTCAGCCCCCTCAGAAAGCTTAAGCAGCATGAGAGCCCGGCCATGCTGGAGCGCTCTCTCAGC 1360
QY 1264 TTGACCCGCTGCTGCTCCAGAGGAGCCAGAGGCGCCCTCTGTTGAGTACCCCAAG 1323
Db 1361 TTGACCCGCTGCTGCTCCAGAGGAGCCAGAGGCGCCCTCTGTTGAGTACCCCAAG 1420
QY 1324 CGAGAGCGATGCTGCTTAATGAGACAGTAGAGAGAGAGAGCTAG-AGATTGAGAGGCT 1382
Db 1421 CGAGAGCGATGCTGCTTAATGAGACAGTAGAGAGAGAGAGCTAGTAGATTGAGAGGCT 1480
QY 1383 TAAGACGAAAGCAAAAGAACTGGAGGCCAAGATGTTGGCCCAAGAGGCTGAGGAAAAGGA 1442
Db 1481 TAAGACGAAAGCAAAAGAACTGGAGGCCAAGATGTTGGCCCAAGAGGCTGAGGAAAAGGA 1540
QY 1443 GAACCATTTGCCCAATG 1461
Db 1541 GAACCATTTGCCCAATG 1559

RESULT 4
US-09-925-300-288
; Sequence 288, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 288
; LENGTH: 2104
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (497)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1323)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-288
Query Match 98.6%; Score 1443.2; DB 9; Length 2104;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1450; Conservative 7; Mismatches 1; Indels 1; Gaps 1;
QY 4 GGTCCCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG 63
Db 67 GGTCCCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG 126
QY 64 GTGGCTGTGCGACTCGGCGCAATTTGTGATGGAAACAGCGGGAGCAAGTATCCCCCTGT 123
Db 127 GTGGCTGTGCGACTCGGCGCAATTTGTGATGGAAACAGCGGGAGCAAGTATCCCCCTGT 186
QY 124 GTGCGGGGATGAGCAGCTGCTCTTAGAGATTGTAACTGTGAGGAAACCAACAGAGACT 183
Db 187 GTGCGGGGATGAGCAGCTGCTCTTAGAGATTGTAACTGTGAGGAAACCAACAGAGACT 246
QY 184 CTCAAATACCAATTTGATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTATGCA 243
Db 247 CTCAAATACCAATTTGATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTATGCA 306
QY 244 GTTCAGTGCAGCCCATCTAAGGCACTTGTGGAAGGCGAGAATGCCAGTGTGCTTGC 303
Db 307 GTTCAGTGCAGCCCATCTAAGGCACTTGTGGAAGGCGAGAATGCCAGTGTGCTTGC 366
QY 304 TATGAACCCACAGGAGCTGGGAAGACGACACAATGTGTGGGAGCCCGAGAGCAACTGGG 363
Db 367 TATGAACCCACAGGAGCTGGGAAGACGACACAATGTGTGGGAGCCCGAGAGCAACTGGG 426
QY 364 GTGATCCCGCGGCTCTCATGACCTCTGAGAGCTCACAAGGGAGAGGTCGCGAGGCG 423
Db 427 GTGATCCCGCGGCTCTCATGACCTCTGAGAGCTCACAAGGGAGAGGTCGCGAGGCG 486
QY 424 CGGCCATGGGCGCTTCTGTCTACCATGCTTACCTAGAGATCTACAGAGAGAGTATTA 483
Db 487 CGGCCATGGGCGCTTCTGTCTACCATGCTTACCTAGAGATCTACAGAGAGAGTATTA 546
QY 484 GACCTCTGTGACCCCTGCTTCCAGAGAGAGCTGCTGTAATCCGAGAGAGAGTCCCGGGGAATATC 543
Db 547 GACCTCTGTGACCCCTGCTTCCAGAGAGAGCTGCTGTAATCCGAGAGAGAGTCCCGGGGAATATC 606
QY 544 CTGATTCGGGCTCTCTCCAGAGAGAGCTGCTGTAATTTGAGTATTTTTCAGCGGCACTTC 603
Db 607 CTGATTCGGGCTCTCTCCAGAGAGAGCTGCTGTAATTTTTCAGCGGCACTTC 666
QY 604 CTGCGCAGCAGTCGAAATCGGAGCTGTAGGAGCCACCGGCTCAACCCAGCGCTCTCCCGC 663
Db 667 CTGCGCAGCAGTCGAAATCGGAGCTGTAGGAGCCACCGGCTCAACCCAGCGCTCTCCCGC 726
QY 664 AGTCATGCTGCTCTGCTGATGAGTGGACGAGGGAACGTTTGGCCCCCATTTGCCAG 723
Db 727 AGTCATGCTGCTCTGCTGATGAGTGGACGAGGGAACGTTTGGCCCCCATTTGCCAG 786
QY 724 CGAGAGGGAAGAACTCTTACCTGATTTGCTTGGCTGAGAGCAACCCGCGCAGCAGGC 783
Db 787 CGAGAGGGAAGAACTCTTACCTGATTTGCTTGGCTGAGAGCAACCCGCGCAGCAGGC 846
QY 784 AACAAAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTGTTGCTCTGGGC 843

Db 847 AACAGGGCCCTTCGGCTAAAGAGAGTGAGGCCATCAACACCTCCCTGTTTGTCTCGGC 906
Qy 844 AAAGTGTAGATGGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGACAGCAAGCTC 903
Db 907 AAAGTGTAGATGGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGACAGCAAGCTC 966
Qy 904 ACTCGCTATTGAGGACTCTCTGGGTGGCTCAGCCCAAGTATCTTTATGCAACATT 963
Db 967 ACTCGCTATTGAGGACTCTCTGGGTGGCTCAGCCCAAGTATCTTTATGCAACATT 1026
Qy 964 GCCCTCAGAGACCTCTACCTAGACACAGCTCCGCACTCAACTTTGCTGCCAGTCC 1023
Db 1027 GCCCTCAGAGACCTCTACCTAGACACAGCTCTCGCACTCAACTTTGCTGCCAGTCC 1086
Qy 1024 AAGGAGGTGATCAATCGGCCCTTTTACCAATGAGAGCCTGAGCCCTCATGCTTGGACCT 1083
Db 1087 AAGGAGGTGATCAATCGGCCCTTTTACCAATGAGAGCCTGAGCCCTCATGCTTGGACCT 1146
Qy 1084 GTTAAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGAGCCCGAGCCCT 1143
Db 1147 GTTAAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGAGCCCGAGCCCT 1206
Qy 1144 GAGAAAGGAGATTTGGAGCCCTGAGCCCATGSCAGCTCCAGCTCTGCTCCAGAAA 1203
Db 1207 GAGAAAGGAGATTTGGAGCCCTGAGCCCATGSCAGCTCCAGCTCTGCTCCAGAAA 1266
Qy 1204 CTGAGCCCTTACAGAAAGCTTAAGCAGCATGAGCCCGCATGCTGGAGCGCTCCT-CAG 1262
Db 1267 CTGAGCCCTTACAGAAAGCTTAAGCAGCATGAGCCCGCATGCTGGAGCGCTCCTCAG 1326
Qy 1263 CTGAGCCCTTACAGAAAGCTTAAGCAGCATGAGCCCGCATGCTGGAGCGCTCCTCAG 1322
Db 1327 CTGAGCCCTTACAGAAAGCTTAAGCAGCATGAGCCCGCATGCTGGAGCGCTCCTCAG 1386
Qy 1323 GCGAGACGGATGTGCTTAATGAAGACAGTAGAAGAGAGCACTGAGATTTGAGAGCT 1382
Db 1387 GCGAGACGGATGTGCTTAATGAAGACAGTAGAAGAGAGCACTGAGATTTGAGAGCT 1446
Qy 1383 TAAGACCAAGCAAAAGAACTGAGAGGCAAGATGCTGGCCAGAGGCTGAGGAAGAGGA 1442
Db 1447 TAAGACCAAGCAAAAGAACTGAGAGGCAAGATGCTGGCCAGAGGCTGAGGAAGAGGA 1506
Qy 1443 GAACCATTTGCCCAATG 1461
Db 1507 GAACCATTTGCCCAATG 1525

RESULT 5
US-10-797-893-5
; Sequence 5, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/797,893
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Human
US-10-797-893-5

Query Match 70.9%; Score 1038.4; DB 17; Length 1041;
Best Local Similarity 99.9%; Pred. No. 9.1e-312;

Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGGGTGCTGTTCGGCTAAAGCAAGATTGGAGCTACTCGTCTGTCACCTCCAGCTCCGCTA 60
Db 1 ATGGGTGCTGTTCGGCTAAAGCAAGATTGGAGCTACTCGTCTGTCACCTCCAGCTCCGCTA 60
Qy 61 AGGTTGCTGTGCGACTTGGCGCCATTGTTGATGGAAACAGCGGAGCAAGTGTATCCCCC 120
Db 61 AGGTTGCTGTGCGACTTGGCGCCATTGTTGATGGAAACAGCGGAGCAAGTGTATCCCCC 120
Qy 121 TGTGTGGGGGCAATGACAGCTCTCTAGAGATTGCTTAAGGAGCAAGTGTATCCCCC 180
Db 121 TGTGTGGGGGCAATGACAGCTCTCTAGAGATTGCTTAAGGAGCAAGTGTATCCCCC 180
Qy 181 ACTCTCAAAATACAGATTGATGCTTCTATGGGAGAGGAGTACTCAGCAGCAATCTAT 240
Db 181 ACTCTCAAAATACAGATTGATGCTTCTATGGGAGAGGAGTACTCAGCAGCAATCTAT 240
Qy 241 GCAGTTTCAAGTGCAGCCCATCTTAAGGCACTTGTGGAAGGCAAGTGTGAGTGTCTT 300
Db 241 GCAGTTTCAAGTGCAGCCCATCTTAAGGCACTTGTGGAAGGCAAGTGTGAGTGTCTT 300
Qy 301 GCCTATGACCCACAGAGCTGGGAGACGACACCAATGCTGGGAGCCCAAGCACT 360
Db 301 GCCTATGACCCACAGAGCTGGGAGACGACACCAATGCTGGGAGCCCAAGCACT 360
Qy 361 GGGTGTATCCCGGGGCTCTCATGGACCTCTCTGAGCTCAAGGAGGAGGCTGCCAG 420
Db 361 GGGTGTATCCCGGGGCTCTCATGGACCTCTCTGAGCTCAAGGAGGAGGCTGCCAG 420
Qy 421 GGGCGGCAATGGGCGCTTCTGTACCATGCTTACCTAGAGATCTACAGAGAGAGTGA 480
Db 421 GGGCGGCAATGGGCGCTTCTGTACCATGCTTACCTAGAGATCTACAGAGAGAGTGA 480
Qy 481 TTAGACCTCTGAGACCTGCTTGGGAGACCTGCTTAATCGAGAGAGTGTGCGGGGAGT 540
Db 481 TTAGACCTCTGAGACCTGCTTGGGAGACCTGCTTAATCGAGAGAGTGTGCGGGGAGT 540
Qy 541 ATCTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGCGGAC 600
Db 541 ATCTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGCGGAC 600
Qy 601 TTCTGCGCAGCCAGTTCGAAATCGGAGCTGTAGGAGCCACCGGCTCAACAGCGCTCTCC 660
Db 601 TTCTGCGCAGCCAGTTCGAAATCGGAGCTGTAGGAGCCACCGGCTCAACAGCGCTCTCC 660
Qy 661 CGCAGTCAATGCTGTGCTCTGCTCAAGGTGGAACAGCGGAAACGTTTGGCGCCATTTGCG 720
Db 661 CGCAGTCAATGCTGTGCTCTGCTCAAGGTGGAACAGCGGAAACGTTTGGCGCCATTTGCG 720
Qy 721 CAGCAGAGGAGAAACTCTACCTGATTGCTTGGCTTAAAGAGTGGAGCCATCAACACCT 780
Db 721 CAGCAGAGGAGAAACTCTACCTGATTGCTTGGCTTAAAGAGTGGAGCCATCAACACCT 780
Qy 781 GGCAACAAGGGCCCTTGGGCTTAAAGAGTGGAGCCATCAACACCTCTCTGTTGCTCGT 840
Db 781 GGCAACAAGGGCCCTTGGGCTTAAAGAGTGGAGCCATCAACACCTCTCTGTTGCTCGT 840
Qy 841 GGCAAAAGTGTAGATGCGCTGAAATCAGGGGCTCTCTGCTGTATCCTTATCGGAGCAGCA 900
Db 841 GGCAAAAGTGTAGATGCGCTGAAATCAGGGGCTCTCTGCTGTATCCTTATCGGAGCAGCA 900
Qy 901 CTCACCTGCTATTGAGGAGTCTCTGGGTGGCTCAGCCCAAGTATCTTATTTGCAAC 960
Db 901 CTCACCTGCTATTGAGGAGTCTCTGGGTGGCTCAGCCCAAGTATCTTATTTGCAAC 960
Qy 961 ATTGCGCCCTGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGG 1020
Db 961 ATTGCGCCCTGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGG 1020
Qy 1021 TCCAGAGGAGTGTATCAATG 1040
Db 1021 TCCAGAGGAGTGTATCAATG 1040

RESULT 6
US-10-797-893-1
; Sequence 1, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/797,893
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1115
; TYPE: DNA
; ORGANISM: Human
US-10-797-893-1

Query Match 70.7%; Score 1035.4; DB 17; Length 1115;
Best Local Similarity 99.9%; Pred. No. 8e-311; Indels 0; Gaps 0;
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	4	GGTCGCTGTCGGCTAAGCAAGATTGGAGTACTCGTCGCTCCACCTCCAGCTCCGCTAAGG	63
DB	78	GGTCGCTGTCGGCTAAGCAAGATTGGAGTACTCGTCGCTCCACCTCCAGCTCCGCTAAGG	137
QY	64	GTGCGCTGTGCGACTCGCGCCATTGTGGATGGAACAGCGGAGCAAGTATCCCCCTGT	123
DB	138	GTGCGCTGTGCGACTCGCGCCATTGTGGATGGAACAGCGGAGCAAGTATCCCCCTGT	197
QY	124	GTGCGGGCATGACAGCTGCTCTTAGAGATTGCTAACTGGAGGACCAACAGGAGACT	183
DB	198	GTGCGGGCATGACAGCTGCTCTTAGAGATTGCTAACTGGAGGACCAACAGGAGACT	257
QY	184	CTCAATATACAGTTTGTATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGCA	243
DB	258	CTCAATATACAGTTTGTATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGCA	317
QY	244	GGTTCAGTCAGGCCATCTTAAGGCATCTTCTGGAAGGGCAGAAATGCTGCTTGCC	303
DB	318	GGTTCAGTCAGGCCATCTTAAGGCATCTTCTGGAAGGGCAGAAATGCTGCTTGCC	377
QY	304	TATGGACCCACAGGAGCTGGGAGAGCGCACAAATGCTGGCAGCCCGAGCAACCTGGG	363
DB	378	TATGGACCCACAGGAGCTGGGAGAGCGCACAAATGCTGGCAGCCCGAGCAACCTGGG	437
QY	364	GTGATCCCGCGGGCTCTCATGGACCTCTCGAGCTCACAGGAGGAGGGTGCGAGGGC	423
DB	438	GTGATCCCGCGGGCTCTCATGGACCTCTCGAGCTCACAGGAGGAGGGTGCGAGGGC	497
QY	424	CGGCATGGGCCCTTTCTGTACCATGTCTTAAGTACTAGAGATCTACAGGAGAGGATTA	483
DB	498	CGGCATGGGCCCTTTCTGTACCATGTCTTAAGTACTAGAGATCTACAGGAGAGGATTA	557
QY	484	GACCTCTGAGACCTTGCTTCGGGAGACCTGGTAAATCCGAGAACTGCGGGGGAAATTC	543
DB	558	GACCTCTGAGACCTTGCTTCGGGAGACCTGGTAAATCCGAGAACTGCGGGGGAAATTC	617
QY	544	CTGATTCGGGGTCTCTCCAGAGACCCATCAGTAGCTTTTCTGATTTTTCAGCGGCCTTC	603
DB	618	CTGATTCGGGGTCTCTCCAGAGACCCATCAGTAGCTTTTCTGATTTTTCAGCGGCCTTC	677
QY	604	CTGCGAGCAGTGAATTCGAGATGTAGAGCCACCGGGCTCAACAGCGCTCTCTCCCGC	663
DB	678	CTGCGAGCAGTGAATTCGAGATGTAGAGCCACCGGGCTCAACAGCGCTCTCTCCCGC	737

QY	664	AGTCATGCTGTGCTCCTGCTCAAGGTGGACCGGAGACGTTTGGCCCCCATTTGCCAG	723
DB	738	AGTCATGCTGTGCTCCTGCTCAAGGTGGACCGGAGACGTTTGGCCCCCATTTGCCAG	797
QY	724	CGAGAGGGGAAAACTCTTACCTGATTGACTTTGGCTGGGTCTAGAGGACAAACCGGCGCAGGC	783
DB	798	CGAGAGGGGAAAACTCTTACCTGATTGACTTTGGCTGGGTCTAGAGGACAAACCGGCGCAGGC	857
QY	784	AACAAGGGCTTCGGCTAAAGAGAGTGGAGCATCAACACCTCTGTTGTCTCTGGGC	843
DB	858	AACAAGGGCTTCGGCTAAAGAGAGTGGAGCATCAACACCTCTGTTGTCTCTGGGC	917
QY	844	AAAGTGGTAGATGCGCTGAATCAGGCGCTCCCTCGGTACCTTATCGGACAGCAAGCTC	903
DB	918	AAAGTGGTAGATGCGCTGAATCAGGCGCTCCCTCGGTACCTTATCGGACAGCAAGCTC	977
QY	904	ACTCGCTTATTGAGGACTCTCTGGGTGGCTCAGCCCAACAGTATCTTATGCAACATT	963
DB	978	ACTCGCTTATTGAGGACTCTCTGGGTGGCTCAGCCCAACAGTATCTTATGCAACATT	1037
QY	964	GCCCTGAGAGACGCTTCTACCTAGACACAGTCTCGGACCTCAACTTGTGCGAGGTCC	1023
DB	1038	GCCCTGAGAGACGCTTCTACCTAGACACAGTCTCGGACCTCAACTTGTGCGAGGTCC	1097
QY	1024	AAGGAGGTGATCAATCG	1040
DB	1098	AAGGAGGTGATCAATCG	1114

RESULT 7
US-09-960-253-131
; Sequence 131, Application US/09960253
; Patent No. US20020123619A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Lodes, Michael J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.556
; CURRENT FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 131
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-960-253-131

Query Match 32.8%; Score 480; DB 9; Length 491;
Best Local Similarity 99.8%; Pred. No. 2.3e-138;
Matches 491; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY	936	AGCCACACAGTATCTTATTTGCCAACATTCGCCCTGAGACGCTTCTACCTAGACACAGT	995
DB	1	AGCCACACAGTATCTTATTTGCCAACATTCGCCCTGAGACGCTTCTACCTAGACACAGT	60
QY	996	CTCCGCACTCAACTTTTGTCTGCGAGGTCCAAGGAGGTGATCAATCGGCCCTTTTACCAATGA	1055
DB	61	CTCCGCACTCAACTTTTGTCTGCGAGGTCCAAGGAGGTGATCAATCGGCCCTTTT-CCAATGA	119
QY	1056	GAGCTGACAGCTCATGCTTGGACCTTAAAGCTGCTCAGAAAGAAATGCTTGGTCC	1115
DB	120	GAGCTGACAGCTCATGCTTGGACCTTAAAGCTGCTCAGAAAGAAATGCTTGGTCC	179
QY	1116	ACCAGAGGCAAGAGAGAGCCGAGGCGCTGAGGAGAGAGATTTGGAGCCCTGAGGCCAT	1175
DB	180	ACCAGAGGCAAGAGAGAGCCGAGGCGCTGAGGAGAGAGATTTGGAGCCCTGAGGCCAT	239
QY	1176	GGAGCTCCAGGCTCTGCTTCCAGAAATCTAGCCCTTACAGAGCTAAGAGCATGGA	1235
DB	240	GGAGCTCCAGGCTCTGCTTCCAGAAATCTAGCCCTTACAGAGCTAAGAGCATGGA	299

Qy 1236 CCGGCCATGTGGAGCCCTCTCAGCTTGGACCGTCTGCTTGCCTCCAGGGGAGCCA 1295
|
Db 300 CCGGCCATGTGGAGCCCTCTCAGCTTGGACCGTCTGCTTGCCTCCAGGGGAGCCA 359
|
Qy 1296 GGGGGCCCTCTGTGTAGTACCCCAAGAGGAGCGGATGTCTAATGAAGACAGTAGA 1355
|
Db 360 GGGGGCCCTCTGTGTAGTACCCCAAGAGGAGCGGATGTCTAATGAAGACAGTAGA 419
|
Qy 1356 AGAAGAGCACCTAGAGATTGAGAGCTTAAAGACGAAGCAAAAGAACTGGAGGCCAAGAT 1415
|
Db 420 AGAAGAGCACCTAGAGATTGAGAGCTTAAAGACGAAGCAAAAGAACTGGAGGCCAAGAT 479
|
Qy 1416 GTTGGCCCGAGAA 1427
|
Db 480 GTTGGCCCGAGAA 491
|

RESULT 8

US-09-803-719-1937

; Sequence 1937, Application US/09803719
; Publication No. US20030044783A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
; APPLICANT: Pot, David
; APPLICANT: Kassam, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, Lee William
; APPLICANT: Stache-Crain, Birgit
; TITLE OF INVENTION: Human Genes and Gene Products
; FILE REFERENCE: 1624.002
; CURRENT APPLICATION NUMBER: US/09/803.719
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/188,609
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 2396
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1937
; LENGTH: 386
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-803-719-1937

Query Match 25.7%; Score 376.4; DB 10; Length 386;
Best Local Similarity 98.4%; Pred. No. 3.3e-106;
Matches 380; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1073 CTTTGGGACTGTAAAGCTGTCTCAGAAAGAAATTTGTTCCACAGAGGCCAAGAGAG 1132
|
Db 1 CGTTGCTGCTCGTTAAAGCTGTCTCAGAAAGAAATTTGTTCCACAGAGGCCAAGAGAG 60
|
Qy 1133 CCGAGGCCCTTAGGAAGAGAGATTGGAGCCCTTAGAGCCCTTAGAGCCATGGCAGCTCCAGCCCTCTG 1192
|
Db 61 CCGAGGCCCTTAGGAAGAGAGATTGGAGCCCTTAGAGCCCTTAGAGCCATGGCAGCTCCAGCCCTCTG 120
|
Qy 1193 CTTCCCAAGAACTCAGCCCTTACAGAAAGCTTAAGACAGCATGGACCCCGCCCTGCTGGAGC 1252
|

Db 121 CTTCCCAAGAACTCAGCCCTTACAGAAAGCTTAAGACAGCATGGACCCCGCCCTGCTGGAGC 180
|
Qy 1253 GCCTCCTCAGCTTGGACCGTCTGCTTCCAGGGGAGCCAGGGGCCCTCTCTGTGA 1312
|
Db 181 GCCTCCTCAGCTTGGACCGTCTGCTTCCAGGGGAGCCAGGGGCCCTCTCTGTGA 240
|
Qy 1313 GTACCCCAAGCGAGAGCGGATGTCTAATGAAGACAGTAGAAGAGACCTTAGAGA 1372
|
Db 241 GTACCCCAAGCGAGAGCGGATGTCTAATGAAGACAGTAGAAGAGACCTTAGAGA 300
|
Qy 1373 TTGAGAGCTTAAGACCAAGCAAAAGAACTGGAGGCCAAGATGTTGGCCCAAGAGGCTG 1432
|
Db 301 TTGAGAGGCTTAAGACCAAGCAAAAGAACTGGAGGCCAAGATGTTGGCCCAAGAGGCTG 360
|
Qy 1433 AGGAAAAGGAGAACCATTTGTCCCA 1458
|
Db 361 AGGAAAAGGAGAACCATTTGTCCCA 386
|

RESULT 9

US-10-087-192-1180
; Sequence 1180, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087.192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1180
; LENGTH: 24525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1180

Query Match 24.9%; Score 365; DB 13; Length 24525;
Best Local Similarity 73.8%; Pred. No. 5.1e-102;
Matches 597; Conservative 0; Mismatches 0; Indels 212; Gaps 2;

Qy 476 AGGTATTAGACCTCTCGACCCCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTGCCGGG 535
|
Db 2313 AGGTATTAGACCTCTCGACCCCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTGCCGGG 2372
|
Qy 536 GGAATATCCTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTTGAGC 595
|
Db 2373 GGAATATCCTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTTGAGC 2432
|
Qy 596 GGCACCTTCCTGCCAGCCAGTCGAAATCGGACTGTAGGAGCACCCGCTCAACACAGCGCT 655
|
Db 2433 GGCACCTTCCTGCCAGCCAGTCGAAATCGGACTGTAGGAGCACCCGCTCAACACAGCGCT 2492
|
Qy 656 CTTCCCGCAGTCATGCTGCTCTCTGTCA----- 685
|
Db 2493 CTTCCCGCAGTCATGCTGCTCTCTGTCAAGGTGAGGCCGCGACAGAGGGCGAGGACCT 2552
|
Qy 686 -----AGGTGACCAAG 696
|
Db 2553 GGGAGCCCGAGGACCTAGCTAAGCACAGACCTTTGTTTACCCCGGAGGACCAAG 2612
|
Qy 697 CGGGAACGTTTGGCCCCCATTTCCGACGAGAGGGAACCTCTACCTGATTGACTTGCT 756
|
Db 2613 CGGGAACGTTTGGCCCCCATTTCCGACGAGAGGGAACCTCTACCTGATTGACTTGCT 2672
|
Qy 757 GGGTCAAGAGCAACCGCGCCACAGGCAACAAGGGCCTTCGGCTAAAGAGAGTGGAGCC 816
|

2673 GGGTCAGAGGACAAACCGGCGCACAGGCAACAGGGGCTTCGGCTAAAAGAGAGTGGAGCC 2732
Db
817 ATCAACACCTCCCTGTTTCTCTGGGCAAGTGGTAGATGCGCTGAATCAGGGCCTCCCT 876
Qy
2733 ATCAACACCTCCCTGTTTCTCTGGGCAAGTGGTAGATGCGCTGAATCAGGGCCTCCCT 2792
Db
877 CGTGACCTTATCGGACAGCAGCTCACTCGCCTATTG----- 915
Qy
2793 CGTGACCTTATCGGACAGCAGCTCACTCGCCTATTG----- 2852
Db
916 ----- 915
Qy
2853 AGGGAAGAGGGGCTCCAGAGAGGTTCTCAGGCTGCTGTTGGGTGGGATAGCAGT 2912
Db
916 -----CAGGACTC 923
Qy
2913 TGAGGCATAGGAAGGCTGGGCTTCTGACCCACCACTGCTGCTCACCCTCAGGACTC 2972
Db
924 TCTGGGTGGCTAGCCACAGTATCTTATGCGCAACATTGCCCTTGAGAGCGCTTCTA 983
Qy
2973 TCTGGGTGGCTAGCCACAGTATCTTATGCGCAACATTGCCCTTGAGAGCGCTTCTA 3032
Db
984 CTTAGACAGCTCTCGCACTCAACTTTGTCGCAAGTCCAGGAGTGAATCAATCGGCC 1043
Qy
3033 CTTAGACAGCTCTCGCACTCAACTTTGTCGCAAGTCCAGGAGTGAATCAATCGGCC 3092
Db
1044 TTTTACCAATGAGAGCTGCGCTCATG 1072
Qy
3093 TTTTACCAATGAGAGCTGCGCTCATG 3121
Db
RESULT 10
US-09-918-995-31705
; Sequence 31705, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31705
; LENGTH: 464
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(464)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-31705
Query Match 18.1%; Score 265; DB 10; Length 464;
Best Local Similarity 77.6%; Pred. No. 1.5e-71; Indels 1; Gaps 1;
Matches 333; Conservative 0; Mismatches 95
Qy 126 CGCGGGATGACAGCTGCTCTCTAGAGATTGCTTAATCGGAGGAACCAACAGGAGACTCT 185
Db 37 CCTTGAATTCATATCGGCTCGATAGAGTGAACCGGCCGAACCAACCGATATCT 96
Qy 186 CAAATACAGTTTATGCTTCTATGAGGAGAGAGTACTCAGCAGACATCTATGAGG 245
Db 97 CCTTACCTGTGCTTTCCTTTTGCAGCCATAGGAGGACTGCGCTGACATGTACCTTCC 156
Qy 246 TTCAGTGCAGCCCTCTTAAGCACTTCTGCGAAGGGCAGAAATGCCAGTGTCTTGCTA 305
Db 157 TTGACCGTGGACCGGTGCGCACTGCGCACTCTCTTTCGTAATTAACGAGGAGCTGGCTA 216
Qy 306 TGGACCCACAGGAGCTGGGAAGACGCAACAATGCTGGGAGCCCAAGCAACCTGGGGT 365

217 CGGTGCTGAGGAGCTGCGATTAC-ATCTCGATGCTTGGCAGCCACAGCAACCTGGGGT 275
Db
366 GATCCCGGGGCTCTCATGACCTCTCGAGCTCACAAGGAGGAGGCTCCGAGGGCCG 425
Qy
276 GATCCCGGGGCTCTCATGACCTCTCGAGCTCACAAGGAGGAGGCTCCGAGGGCCG 335
Db
426 GCCATGGGCTTCTCTGTCACCATGCTTACCTAGAGATCTACCAAGAGAGGTTATTAGA 485
Qy
336 GCCATGGGCTTCTCTGTCACCATGCTTACCTAGAGATCTACCAAGAGAGGTTATTAGA 395
Db
486 CCTCTGTCACCTCTCTGTCGAGAGCCTGTTAATCCGAGAGACTGCGGGGGAATATCCT 545
Qy
396 CCTCTGTCACCTCTCTGTCGAGAGCCTTGGAAATCCGAGAGACTGCGGGGGAATATCCT 455
Db
546 GATTCCGGG 554
Qy
456 GATTCCGGG 464
Db
RESULT 11
US-10-087-192-1177
; Sequence 1177, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1177
; LENGTH: 24291
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(24291)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1177
Query Match 17.9%; Score 262.4; DB 13; Length 24291;
Best Local Similarity 73.8%; Pred. No. 3.9e-70; Indels 76; Gaps 1;
Matches 385; Conservative 0; Mismatches 61
Qy 476 AGGTATTAGACCTCTCGGAGCCTGCTTCGGGAGACTGTTAATCCGAGAGACTGCGGG 535
Db 2939 AGGTATTAGACCTCTTCGGATCTTGGATCTGATCAGGAGACTGTTAATCCGAGAG 2998
Qy 536 GGAATATCTGATTCGGGCTCTCTCCAGAGCCCATAGTAGTCTTGCTGATTTTGGC 595
Db 2999 GGAACATCTGATTCAGGCTCTCACAGAGCCCATCACCAGCTTCTGCTTTCGAGC 3058
Qy 596 GGCATCTTCTGCGAGCAGTGGAAATCGGAGCTGAGGAGCCACCGGCTCAACCGAGCT 655
Db 3059 AGCAGCTTCTTCCAGCCAGTGAATTCGAGCTGAGGAGCCACCGGCTTAACCGAGCT 3118
Qy 656 CCTCCGCGAGTCTGCTGCTGCTCTGCTCA----- 685
Db 3119 CTTCCGAGTCTGAGTCTGAGTCTGCTCTTCTCTCTTCCCTAGGTAGTACAGCT 3178
Qy 686 -----AGTGGACAGCGG 699
Db 3179 AAGCCCTGAAATGTGAGTCTGAGTCTGCTCTTCTCTCTTCCCTAGGTAGTACAGCT 3238
Qy 700 GAACGTTTGGCCCACTTCGCCAGCGAGAGGGGAATACTCTACCTGATTGCTGGCTGGG 759

Db 3239 GAACTTTGACTCCATTTCGCCAGCGGGAAGAAACTCTACCTTATTGATTGGCTGGT 3298
Qy 760 TCAGAGGACAAACCGCGCACAGGCAACAAGGGCCCTTCGGCTAAAGAGAGTGGAGCCATC 819
Db 3299 TCAGAGGACAAACCGTGCACAGGCAACAGGGGCAATTCGGCTCAAGAGAGCGGAGCCATC 3358
Qy 820 AACACCTCCCTGTTTGTCTCGGCAAGAGTGTAGATCGCTTGAATCAAGGGCCCTCCCTCGT 879
Db 3359 AACACCTCCCTGTTTGTACTGGGCAAGTGTGTGATGATTAACACAGGGGCTCCCTCGC 3418
Qy 880 GTACCTTATCGGACAGCAAGCTCACTCGCTATTGAGGAC 921
Db 3419 ATACCATACCGGGACAGCAAGCTCACTCGTCTGTCGAGGTC 3460

RESULT 12

US-10-029-386-10603
; Sequence 10603, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AROMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 10603
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR16.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
; OTHER INFORMATION: NT HIT: AB017334.1, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q14807, EVALUATE 2.00e-36
; OTHER INFORMATION: EST_HUMAN HIT: BF975048.1, EVALUATE 1.00e-126
; OTHER INFORMATION: EST_HUMAN HIT: BF975048.1, EVALUATE 1.00e-126
US-10-029-386-10603

Query Match 15.9%; Score 233.2; DB 15; Length 531;
Best Local Similarity 98.7%; Pred. No. 1.2e-61;
Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 684 CAAGGTGGACCAAGCGGGAACGTTTGGCCCAATTTGCGCCAGCGAGGGAAGAACTCTACCT 743
Db 34 CCAGGTGGACCAAGCGGGAACGTTTGGCCCAATTTGCGCCAGCGAGGGAAGAACTCTACCT 93
Qy 744 GATTGACTTGGCTGGGTTCAGAGGACAAACCGGCGCACAGGCAACAAGGGCCCTTCGGCTAAA 803
Db 94 GATTGACTTGGCTGGGTTCAGAGGACAAACCGGCGCACAGGCAACAAGGGCCCTTCGGCTAAA 153
Qy 804 AGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAGTGTAGATGGCTGAA 863
Db 154 AGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAGTGTAGATGGCTGAA 213
Qy 864 TCAGGGGCTCCCTCGTGTACTTATCGGACAGCAAGCTCACTCGCTATTGAGGAC 921
Db 214 TCAGGGGCTCCCTCGTGTACTTATCGGACAGCAAGCTCACTCGCTATTGAGGAC 271

RESULT 13

US-10-029-386-24303
; Sequence 24303, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AROMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24303
; LENGTH: 232
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR16.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
; OTHER INFORMATION: SWISSPROT HIT: Q14807, EVALUATE 5.00e-37
; OTHER INFORMATION: EST_HUMAN HIT: BE382882.1, EVALUATE 1.00e-125
; OTHER INFORMATION: NT HIT: gi6453817, EVALUATE 1.00e-125
US-10-029-386-24303

Query Match 15.7%; Score 230.4; DB 15; Length 232;
Best Local Similarity 99.6%; Pred. No. 6.4e-61;
Matches 231; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 688 GTGACCAAGCGGGAACGTTTGGCCCAATTTGCGCCAGCGAGGGAAGAACTCTACCTGATT 747
Db 1 GTGACCAAGCGGGAACGTTTGGCCCAATTTGCGCCAGCGAGGGAAGAACTCTACCTGATT 60
Qy 748 GACTTGGCTGGGTTCAGAGGACAAACCGGCGCACAGGCAACAAGGGCCCTTCGGCTAAAAGAG 807
Db 61 GACTTGGCTGGGTTCAGAGGACAAACCGGCGCACAGGCAACAAGGGCCCTTCGGCTAAAAGAG 120
Qy 808 AGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAGTGTGTAGATGGCTGAATCAG 867
Db 121 AGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAGTGTGTAGATGGCTGAATCAG 180
Qy 868 GGCTCCCTCGTGTACTTATCGGACAGCAAGCTCACTCGCTATTGCGAGG 919
Db 181 GGCTCCCTCGTGTACTTATCGGACAGCAAGCTCACTCGCTATTGCGAGG 232

RESULT 14

US-10-108-260A-249
; Sequence 249, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 249
; LENGTH: 2095
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-249

Query Match 12.7%; Score 185.6; DB 16; Length 2095;
Best Local Similarity 53.7%; Pred. No. 1.2e-46;
Matches 464; Conservative 0; Mismatches 379; Indels 21; Gaps 3;
Qy 190 TACCAGTTTGTATGCCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCGAGTTCA 249
Db 286 TACCTGTTCGACGTGGCTTTTGACTTCCCGCCACCCAGGAGATGGTGTATCAGGCCACC 345
Qy 250 GTGCAGCCCATCTTAAGGCACCTTGTGGAAGGCGCAGAAATCCAGTGTGCTTGCCTATGGA 309
Db 346 ACCAAGAGCCCTCATCGAGGGCGTCTATCTCAGGCTACAATGCCACTGTCTTTGCTATGCG 405

QY 310 CCACAGAGAGTGGGAAGACGACACATGCTGGGAGCCCGACAGCAACCTGGGGTGATC 369
 Db |||||
 QY 406 CCACAGAGTGTGGGAACCTACACATGCTGGGACAGACAGGAGCGCTGGCATCTAT 465
 Db |||||
 QY 370 CGCGGGCTCTCATGACCTCTGAGCTCACAGGAGAGGGTGGCGAGGCGCGGCA 429
 Db |||||
 QY 466 GTTCAGACCTCTCAACACCTCTTCCTGCTCATGAGAGACAGCAATGA-----C 516
 Db |||||
 QY 430 TGGGCGCTTTCTGTCAACCATGCTTTACCTAGAGATCTACAGGAGAAGTATTAGACTC 489
 Db |||||
 QY 517 ATGGAGTATGAGTCTCCATGCTTACCTGGAGATCTACATGAGATGATCGGAGACTG 576
 Db |||||
 QY 490 CTGGACCTGCTTCGGGAGACTGTAATCCGAGAAGCTCGGGGGAATCTCTGATT 549
 Db |||||
 QY 577 CTGAACCCCTCTCCCTGGGCTACCTGGAGCTCGGGAGGACTTAAGGGGGGTGATCCAGGTG 636
 Db |||||
 QY 550 CGGGTCTCTCCAGAGCCCATCATGCTTGTCTGATTCTTGGAGCGGCACTTCTTGCCA 609
 Db |||||
 QY 637 GCCGGCATCACGAAGTCTCCACCATCATGCGAAGAGATCATGAGTGTGATGAAG 696
 Db |||||
 QY 610 GCCAGTCGAATCGGACTGTAGAGCCACCGGCTCAACAGAGCTCTCTCCGCGATCAT 669
 Db |||||
 QY 697 GGGACCGGAGAGGACCGAGGAGCCACGGCGCCACAGAGCTCTCTCCGCTCCAC 756
 Db |||||
 QY 670 GCTGTGCTCTGCTCAAGTGGACCGGAGAGCTTTGGCCCATTTGCGCA-----G 723
 Db |||||
 QY 757 GCGGTATCTGAGGTGACCGGCGCCAGCGGAGCGGCTCAAGAACATCTTGCAAGGTG 816
 Db |||||
 QY 724 CGAGAGGAAATCTTACCTGATGCTGCTGGCTCAGAGGACACCGGGCGCACAGCG 783
 Db |||||
 QY 817 CGGAGGGCGCTGTTTCATGCTGAGCTGGCTGGCTCAGAGCGGCTCGCAGACACAG 876
 Db |||||
 QY 784 AACAGGGCTTCGGCTTAAAGAGAGTGGAGCATCAACAGCTCTCTGTTGCTCGGC 843
 Db |||||
 QY 877 AATCGTGGGAGCGTATGAAGAGGGGGGCCACATCAACCGCTCACTGCTGGCACTGGC 936
 Db |||||
 QY 844 AAGTGTGATGCGGTG-----ATCAGGGGCTCTCTGCTGCTTATCGGACAGC 897
 Db |||||
 QY 937 AACTGCATCAACGCGCTCGAGGACAGGGTAGCAACAGTACATCACTATCGGACAGC 996
 Db |||||
 QY 898 AAGCTCACTGCGCTTATGAGGACTCTCTGGGTGGCTCAGCCACAGTATCTTATGTC 957
 Db |||||
 QY 997 AAGCTACCGGCTCTGAGGACTCTCTGGAGGAAACAGCGGCAAGTATCGCT 1056
 Db |||||
 QY 958 AACATTGCGCTGAGAGAGCTTCTTACTAGACACAGTCTCCGCACTCACTTGTGTC 1017
 Db |||||
 QY 1057 CACATGCTCTGCGAGCAGTGTCTTCGAGGAGTCCCGGAACACCCCTGACCTAGCGCGC 1116
 Db |||||
 QY 1018 AGGTCCAAGAGGTGATCAATCGG 1041
 Db |||||
 QY 1117 CGGGCCCAAGACATTAAGACTAGG 1140
 Db |||||

RESULT 15

US-10-287-226-301

; Sequence 301, Application US/10287226

; Publication No. US20040086875A1

; GENERAL INFORMATION:

; APPLICANT: Agee, Michele L.,

; APPLICANT: Alsobrook, John P.,

; APPLICANT: Berghs, Constance,

; APPLICANT: Boldog, Ferenc,

; APPLICANT: Burgess, Catherine E.,

; APPLICANT: Chant, John S.,

; APPLICANT: Chaudhuri, Amitabha,

; APPLICANT: DiPippo, Vincent A.,

; APPLICANT: Eisinger, Shlomit R.,

; APPLICANT: Eisen, Andrew,

; APPLICANT: Ellerman, Karen,

; APPLICANT: Gangolli, Esha A.,

; APPLICANT: Gorman, Linda,

; APPLICANT: Gerlach, Valerie,

; APPLICANT: Ji, Weizhen,

; APPLICANT: Kekuda, Ramesh,
 ; APPLICANT: Khrantsov, Nikolai,
 ; APPLICANT: Li, Li,
 ; APPLICANT: Malyankar, Uriel M.,
 ; APPLICANT: MacDougall, John R.,
 ; APPLICANT: Mezes, Peter S.,
 ; APPLICANT: Millet, Charles E.,
 ; APPLICANT: Millet, Isabelle,
 ; APPLICANT: Oot, Chean Eng,
 ; APPLICANT: Ort, Tatiana,
 ; APPLICANT: Padigaru, Muralidhara,
 ; APPLICANT: Patturajan, Meera,
 ; APPLICANT: Rastelli, Luca,
 ; APPLICANT: Rieser, Daniel K.,
 ; APPLICANT: Rothenberg, Mark E.,
 ; APPLICANT: Shenoy, Suresh G.,
 ; APPLICANT: Spaderna, Steven K.,
 ; APPLICANT: Spatek, Kimberley A.,
 ; APPLICANT: Taupier, Jr., Raymond J.,
 ; APPLICANT: Vernet, Corine A.M.,
 ; APPLICANT: Zerhusen, Bryan D.,
 ; APPLICANT: Zhong, Mei

; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: 21402-480C

; CURRENT APPLICATION NUMBER: US/10/287,226

; CURRENT FILING DATE: 2002-11-04

; PRIOR APPLICATION NUMBER: 60/334,421

; PRIOR FILING DATE: 2001-11-30

; PRIOR APPLICATION NUMBER: 60/354,392

; PRIOR FILING DATE: 2002-02-04

; PRIOR APPLICATION NUMBER: 60/360,148

; PRIOR FILING DATE: 2002-02-27

; PRIOR APPLICATION NUMBER: 60/364,000

; PRIOR FILING DATE: 2002-03-13

; PRIOR APPLICATION NUMBER: 60/404,821

; PRIOR FILING DATE: 2002-08-20

; PRIOR APPLICATION NUMBER: 60/334,526

; PRIOR FILING DATE: 2001-11-30

; PRIOR APPLICATION NUMBER: 60/354,409

; PRIOR FILING DATE: 2002-02-04

; PRIOR APPLICATION NUMBER: 60/364,227

; PRIOR FILING DATE: 2002-03-13

; PRIOR APPLICATION NUMBER: 60/334,027

; PRIOR FILING DATE: 2001-11-28

; PRIOR APPLICATION NUMBER: 60/331,641

; PRIOR FILING DATE: 2001-11-20

; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 673
 ; SOFTWARE: CuraSeqList version 0.1
 ; SEQ ID NO 301
 ; LENGTH: 3374
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (7)..(3351)
 ; US-10-287-226-301

Query Match 10.7%; Score 156.8; DB 16; Length 3374;

Best Local Similarity 54.1%; Pred. No. 1.3e-37;

Matches 395; Conservative 0; Mismatches 317; Indels 18; Gaps 3;

QY 321 TGGGAAGACGACACAAATCTGGGAGCCCGACAGCAACCTGGGGTGATCCCGGGGCTCT 380

Db |||||

QY 612 TGGGAAACCTACACCATCTGGGACAGACAGGAGCTTGTTCAGACCT 671

Db |||||

QY 381 CATGGACCTCTTCAGCTCAGTCCAGAGGAGGTCGCCAGGCGCCCATGGCCCTTTC 440

Db |||||

QY 672 CACAGCACTCTTCGCTGCTCAGAGAGACCAATGA-----CATGGAGTATGA 722

Db |||||

QY 441 TGTCCACATGCTTACCTAGAGATCTTACAGGAGAGGATTTAGACCTCTCGACCTTCG 500

Db |||||

QY 723 GGTCTCATGTCTTACTCTGAGATCTTACATGAGATGATCCCGGAGCTTGTGAACCCCTC 782

Db |||||

Qy 501 TTCGGAGACCTGTGAATCCGAGAGAGTCCGGGGGAATATCCTGATTCGGGTCTCTC 560
 Db |||||
 783 CCTGGCTACCTCGAGCTGCGGAGGACTCTAAGGGGTGATCCAGGTGGCCGCGATCAC 842
 Qy 561 CCAGAAGCCCATAGTAGCTTTGCTGATTTGAGGGGCACTTCCTGCCAGCCAGTCGAAA 620
 Db |||||
 843 CGAAGTCTCCACCATCAATGCAAGGAGATCATGACGCTGCTGATGAAGGGGAACCGGCA 902
 Qy 621 TCGGACTGTAGGAGCCACCCGGCTCAACCCAGCGCTCCTCCCGCAGTCATGCTGTGCTCCT 680
 Db |||||
 903 GAGGACCCAGGAGCCACCGCGCCCAACCCAGACGCTCTCCGCTCCACGCGGTACTGCA 962
 Qy 681 GGTCAAGGTGACCCAGCGGGAACTTTT---GGCCCCATTTCCGCCAGGAGGGGAAAACT 737
 Db |||||
 963 GGTGACCGTGGCCAGCGCGAGCCGGTCAAGAAACATCTTGAGGAGGCGAGGCGCGCCT 1022
 Qy 738 CTACTCATTTGACTTGGTGGTTCAGAGGACAAACCGGGGACAGGCAACAGGCGCTTCG 797
 Db |||||
 1023 GTTCATGATCGACCTGGCTGAGAGCGCGCTCGCAGACACAGAAATCGTGGGACGCG 1082
 Qy 798 GCTAAAAGAGAGTGGAGCCATCAACACCTCCCTGTTGTCCTGGGCAAGTGTAGATGC 857
 Db |||||
 1083 TATGAGGAGGGGSCCAGATCAACCGCTCACTCTGCGACTGGGCAACTGCAATCAAGC 1142
 Qy 858 GCTGAATCAGG-----GCCTCCCTCGTGACCTTATCGGACAGCAAGCTCACTCGCCT 911
 Db |||||
 1143 CCTGAGCGACAAGGTTAGCAACAGTACATCAACTATCGGACAGCAAGCTCACCCGCT 1202
 Qy 912 ATTGACGACTCTCTGGGTGGCTCAGCCGACAGTATCCTTTATGCGCAACATTCGCCCTGA 971
 Db |||||
 1203 CCTGAAGGACTCTCTGGGAGGAAACAGCCGACAGTATGATGCTCAGTCTCAGTCTCTGC 1262
 Qy 972 GAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGGTCCAGGAGGT 1031
 Db |||||
 1263 GAGCAGTCCCTTCAGGAGTCCCGGAACACCTGACCTACGCGCGCGGCCAAGAACAT 1322
 Qy 1032 GATCAATCGG 1041
 Db |||||
 1323 TAAGACTAGG 1332

Search completed: November 11, 2004, 01:20:46
 Job time : 786.387 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 04:57:04 ; Search time 132.833 Seconds
(without alignments)
7833.865 Million cell updates/sec

Title: US-10-797-893-7

Perfect score: 1464

Sequence: 1 atgggtcgctgtcggtgaag.....accattgtccacaatgtga 1464

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1464	100.0	1464	3	US-09-724-224-7
2	1464	100.0	1464	4	US-10-093-317-7
3	1461	99.8	1538	4	US-09-724-224-3
4	1461	99.8	1538	4	US-10-093-317-3
5	1458	99.6	2097	4	US-09-595-684B-34
6	1038.4	70.9	1041	3	US-09-724-224-5
7	1038.4	70.9	1041	4	US-10-093-317-5
8	1035.4	70.7	1115	3	US-09-724-224-1
9	1035.4	70.7	1115	4	US-10-093-317-1
10	138.6	9.5	4108	4	US-09-883-096-1
11	136.6	9.3	1152	4	US-09-883-096-6
12	128.8	8.8	897	3	US-09-621-233-1
13	128.8	8.8	897	3	US-09-724-208-1
14	128.8	8.8	897	3	US-09-724-516-1
15	128.8	8.8	897	4	US-10-090-695-1
16	126.6	8.6	1026	3	US-09-641-806-1
17	126.6	8.6	1026	4	US-09-723-129-1
18	126.6	8.6	1026	4	US-09-722-862-1
19	124.2	8.5	1011	3	US-09-641-806-3
20	124.2	8.5	1011	4	US-09-723-129-3
21	124.2	8.5	1011	4	US-09-722-862-3
22	121.4	8.3	1014	4	US-09-883-096-3
23	101	6.9	3258	4	US-09-799-451-586
24	100	6.8	3572	2	US-08-713-815A-2
25	87.6	6.0	4014	3	US-09-541-782-1
26	87.6	6.0	4014	4	US-09-723-820-1
27	87.6	6.0	4014	4	US-10-270-085-1

28	79.6	5.4	1032	4	US-09-967-908A-7	Sequence 7, Appli
29	79.6	5.4	1032	4	US-10-159-151-7	Sequence 7, Appli
30	79.6	5.4	1122	4	US-09-967-908A-3	Sequence 3, Appli
31	79.6	5.4	1122	4	US-10-159-151-3	Sequence 3, Appli
32	79.4	5.4	3930	3	US-09-162-373-2	Sequence 2, Appli
33	79.4	5.4	3930	3	US-09-467-946-2	Sequence 2, Appli
34	79.2	5.4	1092	4	US-09-967-908A-9	Sequence 9, Appli
35	79.2	5.4	1092	4	US-10-159-151-9	Sequence 9, Appli
36	79.2	5.4	1146	4	US-09-967-908A-5	Sequence 5, Appli
37	79.2	5.4	1146	4	US-10-159-151-5	Sequence 5, Appli
38	79.2	5.4	6409	4	US-09-967-908A-1	Sequence 1, Appli
39	79.2	5.4	6409	4	US-10-159-151-1	Sequence 1, Appli
40	78.4	5.4	1065	3	US-09-724-511-3	Sequence 3, Appli
41	78.4	5.4	1065	4	US-09-723-097-3	Sequence 3, Appli
42	78.4	5.4	1065	4	US-09-632-344-3	Sequence 3, Appli
43	78.4	5.4	1217	3	US-09-724-511-1	Sequence 1, Appli
44	78.4	5.4	1217	4	US-09-723-097-1	Sequence 1, Appli
45	78.4	5.4	1217	4	US-09-632-344-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-724-224-7
; Sequence 7, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09/724, 224
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597, 292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Human
US-09-724-224-7

Query Match		100.0%	Score 1464;	DB 3;	Length 1464;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1464;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGGTGCGTGTGCGGCTAACGAGATTGGAGTACTGCTGCTCCACCTCCAGCTCGGTA	60		
DB	1	ATGGGTGCGTGTGCGGCTAACGAGATTGGAGTACTGCTGCTCCACCTCCAGCTCGGTA	60		
QY	61	AGGGTGGCTGTGCGACTCGCGCCATTGTGGATGGAAACAGCGGAGCAAGTGATCCCCC	120		
DB	61	AGGGTGGCTGTGCGACTCGCGCCATTGTGGATGGAAACAGCGGAGCAAGTGATCCCCC	120		
QY	121	TGTGTGCGGGGCATGGACAGCTGCTCTAGAGATTGCTAACTGGAGAACCAACAGAG	180		
DB	121	TGTGTGCGGGGCATGGACAGCTGCTCTAGAGATTGCTAACTGGAGAACCAACAGAG	180		
QY	181	ACTCTCAATACCAAGTTTGAATGCTTATGCGGAGAGGAGTACTCAGCAGGACATCTAT	240		
DB	181	ACTCTCAATACCAAGTTTGAATGCTTATGCGGAGAGGAGTACTCAGCAGGACATCTAT	240		
QY	241	GCAGGTTCACTGAGGCCCATCTTAAGGCATTTGCTGGAAGGCGAGAAATGCCAGTGTGTT	300		
DB	241	GCAGGTTCACTGAGGCCCATCTTAAGGCATTTGCTGGAAGGCGAGAAATGCCAGTGTGTT	300		
QY	301	GCCTATGGACCCACAGAGCTGGGAAGACGCAACATGTGTGGCAGCCGAGCAACCT	360		
DB	301	GCCTATGGACCCACAGAGCTGGGAAGACGCAACATGTGTGGCAGCCGAGCAACCT	360		
QY	361	GGGGTGTATCCCGGGGCTCTCATGGAGCTCTCATGAGTCTACAGGGAGGAGGTCGCGAG	420		

```

Db 361 GGGGTGATCCCGGGGCTCTCATGGACCTCTCGACCTCACAAAGGAGGAGGTCCGAG 420
QY 421 GCGCGGCCATGGGCCCTTTCTGTACATGTCTTACCTAGAGATCTACAGAGAGGTA 480
Db 421 GCGCGGCCATGGGCCCTTTCTGTACATGTCTTACCTAGAGATCTACAGAGAGGTA 480
QY 481 TTAGACTCTGGACCTCTTCCGAGACCTGTGTAAATCCGAGAGACTCCGGGGGAT 540
Db 481 TTAGACTCTGGACCTCTTCCGAGACCTGTGTAAATCCGAGAGACTCCGGGGGAT 540
QY 541 ATCTGATTCGGGGTCTCTCCAGAGACCATCAGTAGCTTTGCTGATTTGAGCGGCAC 600
Db 541 ATCTGATTCGGGGTCTCTCCAGAGACCATCAGTAGCTTTGCTGATTTGAGCGGCAC 600
QY 601 TTCTGCCAGCCAGTCGAATCGAATCGAATGTAGAGCCACCGGGCTCAACAGCGCTCTCC 660
Db 601 TTCTGCCAGCCAGTCGAATCGAATGTAGAGCCACCGGGCTCAACAGCGCTCTCTCC 660
QY 661 CGCAGTCATGCTGTGCTCTGCTCAAGGTGGAACGAGGGGACGTTTGGCCCATTTCCG 720
Db 661 CGCAGTCATGCTGTGCTCTGCTCAAGGTGGAACGAGGGGACGTTTGGCCCATTTCCG 720
QY 721 CAGCGAGAGGGAATCTCTACCTGATTGACTTGGCTGGGTGAGGACCAACCGGCGACA 780
Db 721 CAGCGAGAGGGAATCTCTACCTGATTGACTTGGCTGGGTGAGGACCAACCGGCGACA 780
QY 781 GGCACAAAGGGCTTCCGGCTTAAAGAGAGTGGAGCCATCAACCTCTCTGTTGCTG 840
Db 781 GGCACAAAGGGCTTCCGGCTTAAAGAGAGTGGAGCCATCAACCTCTCTGTTGCTG 840
QY 841 GGCACAAAGTGTAGATCGGCTGAATCAGGCTCTCCCTCGTGTACTTATCGGAGACAG 900
Db 841 GGCACAAAGTGTAGATCGGCTGAATCAGGCTCTCCCTCGTGTACTTATCGGAGACAG 900
QY 901 CTCACTCGCCTATGTCAGGACTCTCTGGGTGGCTCAGCCACACAGTATCTTATGCGAAC 960
Db 901 CTCACTCGCCTATGTCAGGACTCTCTGGGTGGCTCAGCCACACAGTATCTTATGCGAAC 960
QY 961 ATTGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCGGCTCACTCACTTGTGTCAGG 1020
Db 961 ATTGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCGGCTCACTCACTTGTGTCAGG 1020
QY 1021 TCCAGAGAGTGTATCAATCGGCTTTTACCAATGAGAGCTGAGCTCATGCTTTGGGA 1080
Db 1021 TCCAGAGAGTGTATCAATCGGCTTTTACCAATGAGAGCTGAGCTCATGCTTTGGGA 1080
QY 1081 CTTGTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGCCGAGGC 1140
Db 1081 CTTGTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGCCGAGGC 1140
QY 1141 CTTGAGAGAGAGAGATTGGAGGCTTGAGCCATGGCCATCGAGCTCTGCTCCGAG 1200
Db 1141 CTTGAGAGAGAGAGATTGGAGGCTTGAGCCATGGCCATCGAGCTCTGCTCCGAG 1200
QY 1201 AACTCAGCCCCCTACAGAGCTAAGCAGCATGAGCCCGGCTGCTGGAGGCTCTCTC 1260
Db 1201 AACTCAGCCCCCTACAGAGCTAAGCAGCATGAGCCCGGCTGCTGGAGGCTCTCTC 1260
QY 1261 AGCTTGAGCCCTGCTGCTCCAGGGGAGCCAGGGGCTCTGTTGAGTACCCCA 1320
Db 1261 AGCTTGAGCCCTGCTGCTCCAGGGGAGCCAGGGGCTCTGTTGAGTACCCCA 1320
QY 1321 AAGCGAGAGCGGATGTGCTTAATGAAGACAGTGAAGAGAGAGAGCTTAGAGATTGAGG 1380
Db 1321 AAGCGAGAGCGGATGTGCTTAATGAAGACAGTGAAGAGAGAGAGAGCTTAGAGATTGAGG 1380
QY 1381 CTTAAGACGAGCAAGAAAGAACTGGAGCCAGAGTGTGGCCAGAGGCTGAGGAAAG 1440
Db 1381 CTTAAGACGAGCAAGAAAGAACTGGAGCCAGAGTGTGGCCAGAGGCTGAGGAAAG 1440
QY 1441 GAGAACCATTTGCTCCCAATGTGA 1464

```

```

Db 1441 GAGAACCATTTGCTCCCAATGTGA 1464

RESULT 2
US-10-093-317-7
; Sequence 7, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093,317
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Human
US-10-093-317-7

Query Match 100.0%; Score 1464; DB 4; Length 1464;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1464; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGTGCTGTGCGGTAAAGCAAGATTGGAGCTACTCTGTCTCCACCTCCAGCTCGCGTA 60
Db 1 ATGGGTGCTGTGCGGTAAAGCAAGATTGGAGCTACTCTGTCTCCACCTCCAGCTCGCGTA 60
QY 61 AGGTGGCTGTGCGGTAAAGCAAGATTGGAGCTACTCTGTCTCCACCTCCAGCTCGCGTA 120
Db 61 AGGTGGCTGTGCGGTAAAGCAAGATTGGAGCTACTCTGTCTCCACCTCCAGCTCGCGTA 120
QY 121 TGTGTGCGGGGATGAGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAG 180
Db 121 TGTGTGCGGGGATGAGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAG 180
QY 181 ACTCTCAATATACAGTTTGTATGCTCTTATGGGAGAGAGTACTCTCAGCAGACATCTAT 240
Db 181 ACTCTCAATATACAGTTTGTATGCTCTTATGGGAGAGAGTACTCTCAGCAGACATCTAT 240
QY 241 CAGAGTTTCACTCAGCCCATCTTAAGGCACTTGTCTGGAAGGCGAGAAATGCCAGTGTCT 300
Db 241 CAGAGTTTCACTCAGCCCATCTTAAGGCACTTGTCTGGAAGGCGAGAAATGCCAGTGTCT 300
QY 301 GCCTATGGACCCACAGGAGCTGGGAAGCGCAACAATGCTGGGAGCCGAGCAACCT 360
Db 301 GCCTATGGACCCACAGGAGCTGGGAAGCGCAACAATGCTGGGAGCCGAGCAACCT 360
QY 361 GGGGTGATCCCGGGGCTCTCATGACCTCTTGCAGCTCAAGGGAGAGGTTGCCGAG 420
Db 361 GGGGTGATCCCGGGGCTCTCATGACCTCTTGCAGCTCAAGGGAGAGGTTGCCGAG 420
QY 421 GSCCGGCAATGSGCCCTTTCTGTCAACATGCTTACCTAGAGATCTACAGAGAGAGGTA 480
Db 421 GSCCGGCAATGSGCCCTTTCTGTCAACATGCTTACCTAGAGATCTACAGAGAGAGGTA 480
QY 481 TTAGACTCTCTGGACCTCTTCCGAGAGCCCTGGTAAATCCGAGAGAGCTGCGGGGGAAT 540
Db 481 TTAGACTCTCTGGACCTCTTCCGAGAGCCCTGGTAAATCCGAGAGAGCTGCGGGGGAAT 540
QY 541 ATCTGATTCGGGGTCTCTCCAGAGACCATCAGTAGCTTTGCTGATTTGAGCGGCAC 600
Db 541 ATCTGATTCGGGGTCTCTCCAGAGACCATCAGTAGCTTTGCTGATTTGAGCGGCAC 600
QY 601 TTCTGCCAGCCAGTCGAATCGAATCGAATGTAGAGCCACCGGGCTCAACAGCGCTCTCC 660
Db 601 TTCTGCCAGCCAGTCGAATCGAATGTAGAGCCACCGGGCTCAACAGCGCTCTCTCC 660
QY 661 CGCAGTCATGCTGTGCTCTGCTCAAGGTGGAACGAGGGGACGTTTGGCCCATTTCCG 720

```

Db 661 CGCAGTCATGCTGTGCTCTCTGCTCAAGGTGGACCGAGCGGGAACGTTTGGCCCCCATTTGCG 720
Qy 721 CAGCGAGAGGAGAAACTCTACTGATTCAGTTGGCTGGCTCAGAGGCAACCGGCGACA 780
Db 721 CAGCGAGAGGAGAAACTCTACTGATTCAGTTGGCTGGCTCAGAGGCAACCGGCGACA 780
Qy 781 GCGCAACAGAGGCTTGGCTTAAAGAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTG 840
Db 781 GCGCAACAGAGGCTTGGCTTAAAGAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTG 840
Qy 841 GCGCAACAGAGGCTTGGCTTAAAGAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTG 900
Db 841 GCGCAACAGAGGCTTGGCTTAAAGAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTG 900
Qy 901 CTCACTCGCTTATGTCAGGACTCTCTGGGTGGCTCAGGCCACAGTATCTTTATTGCCAAC 960
Db 901 CTCACTCGCTTATGTCAGGACTCTCTGGGTGGCTCAGGCCACAGTATCTTTATTGCCAAC 960
Qy 961 ATTGCCCTGAGAGCGCTTCTACTAGACAGTCTCCGCACTCAACTTTTGTGCCAGG 1020
Db 961 ATTGCCCTGAGAGCGCTTCTACTAGACAGTCTCCGCACTCAACTTTTGTGCCAGG 1020
Qy 1021 TCCAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTGCAGCTCATGCTTGGGA 1080
Db 1021 TCCAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTGCAGCTCATGCTTGGGA 1080
Qy 1081 CCTGTTAAGCTGTCTCAGAGAGAAATGCTTGGTCCACAGAGGCAAGAGAGCCGAGGC 1140
Db 1081 CCTGTTAAGCTGTCTCAGAGAGAAATGCTTGGTCCACAGAGGCAAGAGAGCCGAGGC 1140
Qy 1141 CCTGAGGAAGAGGAGATTTGGAGCGCTTGGAGCCATGGCAGCTCCAGCTCTGCTTCCAG 1200
Db 1141 CCTGAGGAAGAGGAGATTTGGAGCGCTTGGAGCCATGGCAGCTCCAGCTCTGCTTCCAG 1200
Qy 1201 AAACCTGAGCCCTTACAGAGCTAAGCAGATGAGACCGCGGCATGCTGGAGCGCTCTCT 1260
Db 1201 AAACCTGAGCCCTTACAGAGCTAAGCAGATGAGACCGCGGCATGCTGGAGCGCTCTCT 1260
Qy 1261 AGCTTGAGCGCTGCTGCTTGGCTCCAGGAGGCAAGAGGCGCCCTCTGTTGAGTACCCCA 1320
Db 1261 AGCTTGAGCGCTGCTGCTTGGCTCCAGGAGGCAAGAGGCGCCCTCTGTTGAGTACCCCA 1320
Qy 1321 AAGCGAGAGCGGATGCTGCTAATGAAGACAGTGAAGAGAGGACCTAGAGATTGAGAGG 1380
Db 1321 AAGCGAGAGCGGATGCTGCTAATGAAGACAGTGAAGAGAGGACCTAGAGATTGAGAGG 1380
Qy 1381 CTTAAGACGAGCAAAAGAACTGGAGGCGCAAGATGTTGGCCAGAGAGGCTGAGGAAAAG 1440
Db 1381 CTTAAGACGAGCAAAAGAACTGGAGGCGCAAGATGTTGGCCAGAGAGGCTGAGGAAAAG 1440
Qy 1441 GAGAACCAATTGCTCCCAATGTGA 1464
Db 1441 GAGAACCAATTGCTCCCAATGTGA 1464

RESULT 3

US-09-724-224-3
; Sequence 3, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09/724, 224
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1538

; TYPE: DNA
; ORGANISM: Human
US-09-724-224-3

Query Match 99.8%; Score 1461; DB 3; Length 1538;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GGTGCTGTGCGCTTAAGCAAGATTGGAGCTACTCGTGTGCTCCACCTCCAGCTCGCGTAGG 63
Db 78 GGTGCTGTGCGCTTAAGCAAGATTGGAGCTACTCGTGTGCTCCACCTCCAGCTCGCGTAGG 137
Qy 64 GTGCTGTGCGACTGCGGCCATTTGTGGATGGAACAGCGGAGCAAGTGAATCCCCCTGT 123
Db 138 GTGCTGTGCGACTGCGGCCATTTGTGGATGGAACAGCGGAGCAAGTGAATCCCCCTGT 197
Qy 124 GTGCGGGCATGGACAGCTGCTCTTAGAGATTGCTAACTGGAGGAACACACAGAGACT 183
Db 198 GTGCGGGCATGGACAGCTGCTCTTAGAGATTGCTAACTGGAGGAACACACAGAGACT 257
Qy 184 CTCAAAATACCAATTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 243
Db 258 CTCAAAATACCAATTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 317
Qy 244 GGTTCAGTCAGAGCCCATCTAAGGCACTTGTGGAAGGCGAGAATGCCAGTGTGCTTGGC 303
Db 318 GGTTCAGTCAGAGCCCATCTAAGGCACTTGTGGAAGGCGAGAATGCCAGTGTGCTTGGC 377
Qy 304 TATGGAACCAAGAGAGTGGGAAGACGCAACAATGTGGGACGCCACAGAGCAACCTGGG 363
Db 378 TATGGAACCAAGAGAGTGGGAAGACGCAACAATGTGGGACGCCACAGAGCAACCTGGG 437
Qy 364 GTGATCCCGGGGCTCTCATGGACCTCTGAGAGCTCAAGGAGGAGGGTCCGAGGGC 423
Db 438 GTGATCCCGGGGCTCTCATGGACCTCTGAGAGCTCAAGGAGGAGGGTCCGAGGGC 497
Qy 424 CGGCAATGGGCGCTTCTGTCAACATGCTTTACCTAGAGATCTACACAGGAGAAATATTA 483
Db 498 CGGCAATGGGCGCTTCTGTCAACATGCTTTACCTAGAGATCTACACAGGAGAAATATTA 557
Qy 484 GACCTCTGAGACCTTGTCTGGGAGACCTGGTAAATCCGAGAAGACTGCGGGGGAAATATC 543
Db 558 GACCTCTGAGACCTTGTCTGGGAGACCTGGTAAATCCGAGAAGACTGCGGGGGAAATATC 617
Qy 544 CTGATTCGGGCTCTCTCCAGAGAGCCATCAGTAGCTTGTGATTTTGAAGGCGACCTC 603
Db 618 CTGATTCGGGCTCTCTCCAGAGAGCCATCAGTAGCTTGTGATTTTGAAGGCGACCTC 677
Qy 604 CTGCGAGCCAGTGGAAATCGGACTGTAGGAGCCACCGGCTCAACACAGCGCTCTCCCGC 663
Db 678 CTGCGAGCCAGTGGAAATCGGACTGTAGGAGCCACCGGCTCAACACAGCGCTCTCCCGC 737
Qy 664 AGTCATGCTGTCTCTGCTCAAGGTTGGACAGCGGGAACGTTTGGCCCCCATTTGCGCAG 723
Db 738 AGTCATGCTGTCTCTGCTCAAGGTTGGACAGCGGGAACGTTTGGCCCCCATTTGCGCAG 797
Qy 724 CGAGAGGGAACCTCTACCTGATTGCTTGGCTGGGTTCAGGAGCAACCGGCGACAGGC 783
Db 798 CGAGAGGGAACCTCTACCTGATTGCTTGGCTGGGTTCAGGAGCAACCGGCGACAGGC 857
Qy 784 AACAGGGGCTTCCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGC 843
Db 858 AACAGGGGCTTCCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGC 917
Qy 844 AAAGTGTAGATGCGCTGAATCAGGGCTCCCTGTTGTACCTTATCGGGAACAGAACTC 903
Db 918 AAAGTGTAGATGCGCTGAATCAGGGCTCCCTGTTGTACCTTATCGGGAACAGAACTC 977
Qy 904 ACTGCGCTTATGAGGACTCTCTGGGTGGCTCAGCCACAGTATCTTATTCGCAACT 963
Db 978 ACTGCGCTTATGAGGACTCTCTGGGTGGCTCAGCCACAGTATCTTATTCGCAACT 1037
Qy 964 GCCCTTGAGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCTCAACTTTTGTGCGCAGGTCC 1023


```
Db 1338 TTGGACCGTCTGCTTGGCTCCACAGGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1397
Qy 1324 CGAGAGCGGATGGTCTAATGAAGACAGTGAAGAAGAGGACCTAGAGATTGAGAGGCTT 1383
Db 1398 CGAGAGCGGATGGTCTAATGAAGACAGTGAAGAAGAGGACCTAGAGATTGAGAGGCTT 1457
Qy 1384 AAGACGAAGCAAAAAGAACTGGAGGCCCAAGATGTTGGCCCCAGAAAGGCTGAGGAAAAGGAG 1443
Db 1458 AAGACGAAGCAAAAAGAACTGGAGGCCCAAGATGTTGGCCCCAGAAAGGCTGAGGAAAAGGAG 1517
Qy 1444 AACCAATTGTCCCAATGTGA 1464
Db 1518 AACCAATTGTCCCAATGTGA 1538

RESULT 5
US-09-595-684B-34
; Sequence 34, Application US/09595684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Ohashi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; FILE REFERENCE: cytop036
; CURRENT APPLICATION NUMBER: US/09/595,684B
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 2097
; TYPE: DNA
; ORGANISM: Human
US-09-595-684B-34

Query Match 99.6%; Score 1458; DB 4; Length 2097;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GGTGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGGTAAGG 63
Db 100 GGTGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGGTAAGG 159
Qy 64 GTGGCTGTGGACTGCGGCCATTTGTGGATGGAAACAGCGGAGCAAGTGTATCCCTGT 123
Db 160 GTGGCTGTGGACTGCGGCCATTTGTGGATGGAAACAGCGGAGCAAGTGTATCCCTGT 219
Qy 124 GTGGCGGCGATGGAACAGCTGCTCTAGAGATTGCTAAGTGGAGAAACCAACAGGAGACT 183
Db 220 GTGGCGGCGATGGAACAGCTGCTCTAGAGATTGCTAAGTGGAGAAACCAACAGGAGACT 279
Qy 184 CTCAAATACCAAGTTGATGCTTCTATGGGAGAGGAGTACTACGAGGACATCTATGCA 243
Db 280 CTCAAATACCAAGTTGATGCTTCTATGGGAGAGGAGTACTACGAGGACATCTATGCA 339
Qy 244 GGTTCAGTCAGCGCCATCTTAAGCACTTGTGGAAGGCGCAAGTGTGCTGCTTGGC 303
Db 340 GGTTCAGTCAGCGCCATCTTAAGCACTTGTGGAAGGCGCAAGTGTGCTGCTTGGC 399
Qy 304 TATGACCCCAAGAGCTGGGAAGACGCAACAATGCTGGCGAGCCCGAGAGCAACCTGGG 363
Db 400 TATGACCCCAAGAGCTGGGAAGACGCAACAATGCTGGCGAGCCCGAGAGCAACCTGGG 459
Qy 364 GTGATCCCGGGGCTCTCATGGACCTCTCGAGCTCAAGAGGAGGAGGCTGCGGAGGC 423
Db 460 GTGATCCCGGGGCTCTCATGGACCTCTCGAGCTCAAGAGGAGGAGGCTGCGGAGGC 519
```

```
Qy 424 CGGCATGGGGCCCTTTCTGTCAACATGTCTTACCTAGAGATCTACACGAGAGAGGTATTA 483
Db 520 CGGCATGGGGCCCTTTCTGTCAACATGTCTTACCTAGAGATCTACACGAGAGAGGTATTA 579
Qy 484 GACCTCTCGGACCCCTGCTTCCGGAGAGACTGTGTAATCCGAGAGAGACTGCGGGGGGAATATC 543
Db 580 GACCTCTCGGACCCCTGCTTCCGGAGAGACTGTGTAATCCGAGAGAGACTGCGGGGGGAATATC 639
Qy 544 CTGATTCGGGGTCTCTCCCAAGCCCATCAGTAGCTTTGCTGATTTTGGGGGACATTC 603
Db 640 CTGATTCGGGGTCTCTCCCAAGCCCATCAGTAGCTTTGCTGATTTTGGGGGACATTC 699
Qy 604 CTGCCAGCCAGTTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACGAGGCTCTCTCCGCG 663
Db 700 CTGCCAGCCAGTTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACGAGGCTCTCTCCGCG 759
Qy 664 AGTCATGCTGTCTCTGCTCAAGGTGGAACAGCGGGGAAACGTTTGGGCCCATTTGCGCAG 723
Db 760 AGTCATGCTGTCTCTGCTCAAGGTGGAACAGCGGGGAAACGTTTGGGCCCATTTGCGCAG 819
Qy 724 CGAGAGGGAAAACCTACCTGATTTGACTTGGCTGGGTCAAGAGGACAAACCGGGGACAGGC 783
Db 820 CGAGAGGGAAAACCTACCTGATTTGACTTGGCTGGGTCAAGAGGACAAACCGGGGACAGGC 879
Qy 784 AACAAAGGGCCCTTCGGCTAAAAGAGAGTGGAGCCCATCAACACCTCCCTGTTTGTCTGGGC 843
Db 880 AACAAAGGGCCCTTCGGCTAAAAGAGAGTGGAGCCCATCAACACCTCCCTGTTTGTCTGGGC 939
Qy 844 AAAGTGTAGATGCGCTGAATCAGGGGCTCTCTGCTGTACCTTATCGGGAACAGCAAGCTC 903
Db 940 AAAGTGTAGATGCGCTGAATCAGGGGCTCTCTGCTGTACCTTATCGGGAACAGCAAGCTC 999
Qy 904 ACTCGCTATTCAGAGGACTCTCTGGGTGGCTCAGCCCAACATATCTTATTTGCCACATTT 963
Db 1000 ACTCGCTATTCAGAGGACTCTCTGGGTGGCTCAGCCCAACATATCTTATTTGCCACATTT 1059
Qy 964 GCCCTGAGAGAGCGCTTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGGTGCC 1023
Db 1060 GCCCTGAGAGAGCGCTTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGGTGCC 1119
Qy 1024 AAGAGGTGATCAATCGGCTTTTACCAATGAGAGCTCGAGCTCATGCTTGGGACCT 1083
Db 1120 AAGAGGTGATCAATCGGCTTTTACCAATGAGAGCTCGAGCTCATGCTTGGGACCT 1179
Qy 1084 GTTAAGCTGTCTCAGAAAGAAATTGCTTGGTCCACAGAGCAAGAGAGCCGAGGCGCT 1143
Db 1180 GTTAAGCTGTCTCAGAAAGAAATTGCTTGGTCCACAGAGCAAGAGAGCCGAGGCGCT 1239
Qy 1144 GAGGAAGAGAGAGATTGGGAGCCCTGAGCCCATGSCAGCTCCAGGCTCTGCTCCAGAAA 1203
Db 1240 GAGGAAGAGAGAGATTGGGAGCCCTGAGCCCATGSCAGCTCCAGGCTCTGCTCCAGAAA 1299
Qy 1204 CTGAGCCCCCTCAGAAAGCTAAGCAGCATGGAACCCGGCCATGTGGAGCGCTCTCTCAGC 1263
Db 1300 CTGAGCCCCCTCAGAAAGCTAAGCAGCATGGAACCCGGCCATGTGGAGCGCTCTCTCAGC 1359
Qy 1264 TTGAGCGCTGTCTGCTCCAGGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1323
Db 1360 TTGAGCGCTGTCTGCTCCAGGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1419
Qy 1324 CGAGAGCGGATGGTGTCTAATGAAGACAGTGAAGAAGAGAGGACCTTAGAGATTGAGAGGCTT 1383
Db 1420 CGAGAGCGGATGGTGTCTAATGAAGACAGTGAAGAAGAGAGGACCTTAGAGATTGAGAGGCTT 1479
Qy 1384 AAGACGAAGCAAAAAGAACTGGAGGCCCAAGATGTTGGCCCCAGAAAGGCTGAGGAAAAGGAG 1443
Db 1480 AAGACGAAGCAAAAAGAACTGGAGGCCCAAGATGTTGGCCCCAGAAAGGCTGAGGAAAAGGAG 1539
Qy 1444 AACCAATTGTCCCAATG 1461
Db 1540 AACCAATTGTCCCAATG 1557
```

```

RESULT 6
US-09-724-224-5
; Sequence 5, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09/724,224
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Human
US-09-724-224-5

Query Match          70.9%; Score 1038.4; DB 3; Length 1041;
Best Local Similarity 99.9%; Pred. No. 8.8e-293;
Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGTGCGTGTGCGGCTAAGCAAGATTGGAGTCTACTCGTCCACCTCCAGCTCGCGTA 60
DB 1 ATGGGTGCGTGTGCGGCTAAGCAAGATTGGAGTCTACTCGTCCACCTCCAGCTCGCGTA 60
QY 61 AGGGTGGCTGTGCGACTCGGGCCATTGTGGATGGAACAGCGGGAGCAAGTATCCCCC 120
DB 61 AGGGTGGCTGTGCGACTCGGGCCATTGTGGATGGAACAGCGGGAGCAAGTATCCCCC 120
QY 121 TGTGTGCGGGCATGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACACCAAGAG 180
DB 121 TGTGTGCGGGCATGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACACCAAGAG 180
QY 121 TGTGTGCGGGCATGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACACCAAGAG 180
DB 121 TGTGTGCGGGCATGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACACCAAGAG 180
QY 181 ACTCTCAAAATACCAAGTTTGTATGCTTCTATGGGGAGAGAGTACTCAGCAGCATCTAT 240
DB 181 ACTCTCAAAATACCAAGTTTGTATGCTTCTATGGGGAGAGAGTACTCAGCAGCATCTAT 240
QY 181 ACTCTCAAAATACCAAGTTTGTATGCTTCTATGGGGAGAGAGTACTCAGCAGCATCTAT 240
DB 181 ACTCTCAAAATACCAAGTTTGTATGCTTCTATGGGGAGAGAGTACTCAGCAGCATCTAT 240
QY 241 GCAGGTTCAAGTGCAGGCCCATCTCCTAGGCACTTGTCTGGAAGGGCAGATGCCAGTGTCT 300
DB 241 GCAGGTTCAAGTGCAGGCCCATCTCCTAGGCACTTGTCTGGAAGGGCAGATGCCAGTGTCT 300
QY 241 GCAGGTTCAAGTGCAGGCCCATCTCCTAGGCACTTGTCTGGAAGGGCAGATGCCAGTGTCT 300
DB 241 GCAGGTTCAAGTGCAGGCCCATCTCCTAGGCACTTGTCTGGAAGGGCAGATGCCAGTGTCT 300
QY 301 GCCTATGAGCCCAAGAGCTGGGAAGAGCGACAAATGCTGGGAGCCCGCAGAACCT 360
DB 301 GCCTATGAGCCCAAGAGCTGGGAAGAGCGACAAATGCTGGGAGCCCGCAGAACCT 360
QY 361 GGGGTGATCCCGGGGCTCTCATGGAACCTCTGAGCTCAAGAGGAGAGGGTGGCGAG 420
DB 361 GGGGTGATCCCGGGGCTCTCATGGAACCTCTGAGCTCAAGAGGAGAGGGTGGCGAG 420
QY 421 GGGGGGCAATGGGCGCTTCTGTCACCAATGCTTACCTAGAGATCTACAGGAGAGAGTA 480
DB 421 GGGGGGCAATGGGCGCTTCTGTCACCAATGCTTACCTAGAGATCTACAGGAGAGAGTA 480
QY 481 TTAGACCTCTCTGAGACCTCTGCTTCCGAGAGACCTGCTAAATCCGAGAGACTCCCGGGGAAT 540
DB 481 TTAGACCTCTCTGAGACCTCTGCTTCCGAGAGACCTGCTAAATCCGAGAGACTCCCGGGGAAT 540
QY 541 ATCTGATTCGGGTCTCTCCAGAGCCCATCATGAGTCTGCTGATTTTGGCGGCAC 600
DB 541 ATCTGATTCGGGTCTCTCCAGAGCCCATCATGAGTCTGCTGATTTTGGCGGCAC 600
QY 601 TTCTCTGCCAGCAAGTCCGAATTCGACTGTAGAGCCACCCGGCTCAACAGCGCTCTCTCC 660
DB 601 TTCTCTGCCAGCAAGTCCGAATTCGACTGTAGAGCCACCCGGCTCAACAGCGCTCTCTCC 660
QY 661 CGCAGTCAATGCTGTCTCTCTGCTCAAGGTGAGCAAGCGGGAAACGTTTGGGCCCAATTCGC 720
DB 661 CGCAGTCAATGCTGTCTCTCTGCTCAAGGTGAGCAAGCGGGAAACGTTTGGGCCCAATTCGC 720

RESULT 7
US-10-093-317-5
; Sequence 5, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093,317
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Human
US-10-093-317-5

Query Match          70.9%; Score 1038.4; DB 4; Length 1041;
Best Local Similarity 99.9%; Pred. No. 8.8e-293;
Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGTGCGTGTGCGGCTAAGCAAGATTGGAGTCTACTCGTCCACCTCCAGCTCGCGTA 60
DB 1 ATGGGTGCGTGTGCGGCTAAGCAAGATTGGAGTCTACTCGTCCACCTCCAGCTCGCGTA 60
QY 61 AGGGTGGCTGTGCGACTCGGGCCATTGTGGATGGAACAGCGGGAGCAAGTATCCCCC 120
DB 61 AGGGTGGCTGTGCGACTCGGGCCATTGTGGATGGAACAGCGGGAGCAAGTATCCCCC 120
QY 121 TGTGTGCGGGCATGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACACCAAGAG 180
DB 121 TGTGTGCGGGCATGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACACCAAGAG 180
QY 121 TGTGTGCGGGCATGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACACCAAGAG 180
DB 121 TGTGTGCGGGCATGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACACCAAGAG 180
QY 181 ACTCTCAAAATACCAAGTTTGTATGCTTCTATGGGGAGAGAGTACTCAGCAGCATCTAT 240
DB 181 ACTCTCAAAATACCAAGTTTGTATGCTTCTATGGGGAGAGAGTACTCAGCAGCATCTAT 240
QY 181 ACTCTCAAAATACCAAGTTTGTATGCTTCTATGGGGAGAGAGTACTCAGCAGCATCTAT 240
DB 181 ACTCTCAAAATACCAAGTTTGTATGCTTCTATGGGGAGAGAGTACTCAGCAGCATCTAT 240
QY 241 GCAGGTTCAAGTGCAGGCCCATCTCCTAGGCACTTGTCTGGAAGGGCAGATGCCAGTGTCT 300
DB 241 GCAGGTTCAAGTGCAGGCCCATCTCCTAGGCACTTGTCTGGAAGGGCAGATGCCAGTGTCT 300
QY 241 GCAGGTTCAAGTGCAGGCCCATCTCCTAGGCACTTGTCTGGAAGGGCAGATGCCAGTGTCT 300
DB 241 GCAGGTTCAAGTGCAGGCCCATCTCCTAGGCACTTGTCTGGAAGGGCAGATGCCAGTGTCT 300
QY 301 GCCTATGAGCCCAAGAGCTGGGAAGAGCGACAAATGCTGGGAGCCCGCAGAACCT 360
DB 301 GCCTATGAGCCCAAGAGCTGGGAAGAGCGACAAATGCTGGGAGCCCGCAGAACCT 360
QY 361 GGGGTGATCCCGGGGCTCTCATGGAACCTCTGAGCTCAAGAGGAGAGGGTGGCGAG 420
DB 361 GGGGTGATCCCGGGGCTCTCATGGAACCTCTGAGCTCAAGAGGAGAGGGTGGCGAG 420
QY 421 GGGGGGCAATGGGCGCTTCTGTCACCAATGCTTACCTAGAGATCTACAGGAGAGAGTA 480
DB 421 GGGGGGCAATGGGCGCTTCTGTCACCAATGCTTACCTAGAGATCTACAGGAGAGAGTA 480
QY 481 TTAGACCTCTCTGAGACCTCTGCTTCCGAGAGACCTGCTAAATCCGAGAGACTCCCGGGGAAT 540
DB 481 TTAGACCTCTCTGAGACCTCTGCTTCCGAGAGACCTGCTAAATCCGAGAGACTCCCGGGGAAT 540
QY 541 ATCTGATTCGGGTCTCTCCAGAGCCCATCATGAGTCTGCTGATTTTGGCGGCAC 600
DB 541 ATCTGATTCGGGTCTCTCCAGAGCCCATCATGAGTCTGCTGATTTTGGCGGCAC 600
QY 601 TTCTCTGCCAGCAAGTCCGAATTCGACTGTAGAGCCACCCGGCTCAACAGCGCTCTCTCC 660
DB 601 TTCTCTGCCAGCAAGTCCGAATTCGACTGTAGAGCCACCCGGCTCAACAGCGCTCTCTCC 660
QY 661 CGCAGTCAATGCTGTCTCTCTGCTCAAGGTGAGCAAGCGGGAAACGTTTGGGCCCAATTCGC 720
DB 661 CGCAGTCAATGCTGTCTCTCTGCTCAAGGTGAGCAAGCGGGAAACGTTTGGGCCCAATTCGC 720

```


Db 301 GCCTATGACCAACGAGAGCTGGGAGAGCGCACAAATGCTGGGAGCCCGCAGAGCAACCT 360
 Qy 361 GGGGTGATCCGCGGGCTCTCATGGACCTCTGAGACCTCTGAGCTCAACAAGGGAGAGGGTCCGAG 420
 Db 361 GGGGTGATCCGCGGGCTCTCATGGACCTCTGAGACCTCTGAGCTCAACAAGGGAGAGGGTCCGAG 420
 Qy 421 GCGCGGCATGGGCGCTTCTGTACCATGCTTACCTAGAGATCTACAGAGAGGTA 480
 Db 421 GCGCGGCATGGGCGCTTCTGTACCATGCTTACCTAGAGATCTACAGAGAGGTA 480
 Qy 481 TTAGACCTCTCGGACCTCTCGGAGACCTGCTGTAATCCGAGAGACTGCGCGGGGAAT 540
 Db 481 TTAGACCTCTCGGACCTCTCGGAGACCTGCTGTAATCCGAGAGACTGCGCGGGGAAT 540
 Qy 541 ATCTGATTCGGGCTCTCTCCAGAAAGCCCTACAGTACCTTGTGATTTGAGCGGCAC 600
 Db 541 ATCTGATTCGGGCTCTCTCCAGAAAGCCCTACAGTACCTTGTGATTTGAGCGGCAC 600
 Qy 601 TTCTGCGCAGCAGTGGAAATCGGACTGTAGAGGCAACCGGCTCAACAGCGTCTCTCC 660
 Db 601 TTCTGCGCAGCAGTGGAAATCGGACTGTAGAGGCAACCGGCTCAACAGCGTCTCTCC 660
 Qy 661 GCGAGTCAATGCTGCTCTCTGCTCAAGTGGACAGCGGGAACGTTTGGCCCAATTCGC 720
 Db 661 GCGAGTCAATGCTGCTCTCTGCTCAAGTGGACAGCGGGAACGTTTGGCCCAATTCGC 720
 Qy 721 CAGCGAGAGGAAATCTACTGATTTGACTTGGCTGGGTCAAGGCAACCGCGCAC 780
 Db 721 CAGCGAGAGGAAATCTACTGATTTGACTTGGCTGGGTCAAGGCAACCGCGCAC 780
 Qy 781 GCGAAAGGCGCTTGGCTTAAAGAGAGTGGGCAATCAACCTCCCTGTTTGTCTG 840
 Db 781 GCGAAAGGCGCTTGGCTTAAAGAGAGTGGGCAATCAACCTCCCTGTTTGTCTG 840
 Qy 841 GCGAAAGTGTAGATGCGCTGAATCAGGCGCTCCCTGCTGCTTATCGGAGCAG 900
 Db 841 GCGAAAGTGTAGATGCGCTGAATCAGGCGCTCCCTGCTGCTTATCGGAGCAG 900
 Qy 901 CTCACTCGCTTATGAGGACTCTCTGGGTGGCTCAGCGCCACAGTATCTTATGCGCAAC 960
 Db 901 CTCACTCGCTTATGAGGACTCTCTGGGTGGCTCAGCGCCACAGTATCTTATGCGCAAC 960
 Qy 961 ATTGCGCTGAGAGACGCTTCTACTAGACACAGTCTCGGCACTCAACTTTGTCGCGAG 1020
 Db 961 ATTGCGCTGAGAGACGCTTCTACTAGACACAGTCTCGGCACTCAACTTTGTCGCGAG 1020
 Qy 1021 TCCAAGGAGGTGATCAATCG 1040
 Db 1021 TCCAAGGAGGTGATCAATTG 1040

RESULT 8

US-09-724-224-1
 ; Sequence 1, Application US/09724224
 ; Patent No. 6387644
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
 ; their use
 ; FILE REFERENCE: 1044
 ; CURRENT APPLICATION NUMBER: US/09/724,224
 ; CURRENT FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: 09/597,292
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1115
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-724-224-1

Query Match

70.7%; Score 1035.4; DB 3; Length 1115;

Best Local Similarity 99.9%; Pred. No. 6.9e-292;
 Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 4 GGTGCTGTGGCTTAAGCAAGATTGGAGCTACTCGTCTCACTCCAGCTCGCGTAAGG 63
 Db 78 GGTGCTGTGGCTTAAGCAAGATTGGAGCTACTCGTCTCACTCCAGCTCGCGTAAGG 137
 Qy 64 GTGGCTGTGGCTTAAGCAAGATTGGAGCTACTCGTCTCACTCCAGCTCGCGTAAGG 123
 Db 138 GTGGCTGTGGCTTAAGCAAGATTGGAGCTACTCGTCTCACTCCAGCTCGCGTAAGG 197
 Qy 124 GTGGGCGCATGAGACAGCTGCTCTAGAGATGCTAACTGGAGAAACCAAGAGACT 183
 Db 198 GTGGGCGCATGAGACAGCTGCTCTAGAGATGCTAACTGGAGAAACCAAGAGACT 257
 Qy 184 CTCAATATACAGATTGATGCTCTATGGGGAGAGAGTACTCAGCAGGACATCTATGCA 243
 Db 258 CTCAATATACAGATTGATGCTCTATGGGGAGAGAGTACTCAGCAGGACATCTATGCA 317
 Qy 244 GGTTCAGTCAGCAGCCATCTTAAGGCACTTGTGGAAGGGCAGAAATGCCAGTGTCTGCC 303
 Db 318 GGTTCAGTCAGCAGCCATCTTAAGGCACTTGTGGAAGGGCAGAAATGCCAGTGTCTGCC 377
 Qy 304 TATGGAACCAAGAGCTGGGAGAGCAACAATGTGGGCAACCAAGAGAAACCTTGG 363
 Db 378 TATGGAACCAAGAGCTGGGAGAGCAACAATGTGGGCAACCAAGAGAAACCTTGG 437
 Qy 364 GTGATCCGCGGGCTCTCATGGACCTCTGAGGAGCTCAAGGGAGAGGGTCCGAGGGC 423
 Db 438 GTGATCCGCGGGCTCTCATGGACCTCTGAGGAGCTCAAGGGAGAGGGTCCGAGGGC 497
 Qy 424 CGGCCATGGGCTCTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGAAATGTA 483
 Db 498 CGGCCATGGGCTCTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGAAATGTA 557
 Qy 484 GACCTCTGAGACCTGCTTCGGGAGAGCTGGTAAATCGGAGAAATATC 543
 Db 558 GACCTCTGAGACCTGCTTCGGGAGAGCTGGTAAATCGGAGAAATATC 617
 Qy 544 CTGATTCGGGCTCTTCCAGAGAGCCATCAGTAGCTTGTGATTTTGGCGGCACTTC 603
 Db 618 CTGATTCGGGCTCTTCCAGAGAGCCATCAGTAGCTTGTGATTTTGGCGGCACTTC 677
 Qy 604 CTGCGACCACTGAAATCGGAGCTGTAGGAGCACCAGGCTCAACAGCGCTCTCCGCG 663
 Db 678 CTGCGACCACTGAAATCGGAGCTGTAGGAGCACCAGGCTCAACAGCGCTCTCCGCG 737
 Qy 664 AGTCATGCTGTCTCTGGTCAAGGAGAGCAAGGGGAAAGCTTTGGGCGGCAATTCGCG 723
 Db 738 AGTCATGCTGTCTCTGGTCAAGGAGAGCAAGGGGAAAGCTTTGGGCGGCAATTCGCG 797
 Qy 724 CGAGAGGAGAACTCTACCTGATTTGACTTGGCTGGGTCAAGGAGCAACCGGCGCAGGC 783
 Db 798 CGAGAGGAGAACTCTACCTGATTTGACTTGGCTGGGTCAAGGAGCAACCGGCGCAGGC 857
 Qy 784 AACAGGGGCTCTCGGCTTAAAGAGAGTGGAGCCATCAACCTCTCTGTTTGTCTGGCG 843
 Db 858 AACAGGGGCTCTCGGCTTAAAGAGAGTGGAGCCATCAACCTCTCTGTTTGTCTGGCG 917
 Qy 844 AAAGTGTAGATGCGTGAATCAGGGCTCTCTGCTGCTTATTCGGGAGCAGCAAGCTC 903
 Db 918 AAAGTGTAGATGCGTGAATCAGGGCTCTCTGCTGCTTATTCGGGAGCAGCAAGCTC 977
 Qy 904 ACTGCGCTTATGAGGACTCTCTGGGTGGCTCAGCCACAGATATCTTATTCGCAACT 963
 Db 978 ACTGCGCTTATGAGGACTCTCTGGGTGGCTCAGCCACAGATATCTTATTCGCAACT 1037
 Qy 964 GCCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGAGCTCAACTTTGTCGCGAGTCC 1023
 Db 1038 GCCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGAGCTCAACTTTGTCGCGAGTCC 1097
 Qy 1024 AAGGAGGTGATCAATCG 1040


```
Db 1098 AAGGAGGTGATCAATTG 1114
RESULT 9
US-10-093-317-1
; Sequence 1, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1115
; TYPE: DNA
; ORGANISM: Human
US-10-093-317-1
Query Match 70.7%; Score 1035.4; DB 4; Length 1115;
Best Local Similarity 99.9%; Pred. No. 6.9e-292;
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 GGTGCTGTCGGCTAAGCAAGATTGGAGCTACTGCTGCCACCTCCAGCTCGCGTAAGG 63
Db 78 GGTGCTGTCGGCTAAGCAAGATTGGAGCTACTGCTGCCACCTCCAGCTCGCGTAAGG 137
QY 64 GTGCTGTGCGACTCGCGGCAATTTGGATGGAACAGCGGAGCAAGTATCCCCCTGT 123
Db 138 GTGCTGTGCGACTCGCGGCAATTTGGATGGAACAGCGGAGCAAGTATCCCCCTGT 197
QY 124 GTGCGGGGATGACAGCTGCTCTAGAGATTGCTAACTGGAGGACCAACAGGAGACT 183
Db 198 GTGCGGGGATGACAGCTGCTCTAGAGATTGCTAACTGGAGGACCAACAGGAGACT 257
QY 184 CTCAATACAGTTTGTATGCTTCTATGGGAGGAGTACTCAGCAGGACATCTATGCA 243
Db 258 CTCAATACAGTTTGTATGCTTCTATGGGAGGAGTACTCAGCAGGACATCTATGCA 317
QY 244 GGTTCAGTCAGGCCATCTTAAGCACTTGTGGAAGGCGAGATGCGAGTGTGCTGCC 303
Db 318 GGTTCAGTCAGGCCATCTTAAGCACTTGTGGAAGGCGAGATGCGAGTGTGCTGCC 377
QY 304 TATGACCCCAAGGAGCTGGAGACGACACACAAATGCTGGGAGCCCGAGCAACCTGGG 363
Db 378 TATGACCCCAAGGAGCTGGAGACGACACACAAATGCTGGGAGCCCGAGCAACCTGGG 437
QY 364 GTGATCCCGGGGCTCTCATGGAACCTCTGAGCTCAAGGGAGGAGGTGCGAGGGC 423
Db 438 GTGATCCCGGGGCTCTCATGGAACCTCTGAGCTCAAGGGAGGAGGTGCGAGGGC 497
QY 424 CGGCAATGGGCGCTTCTGTACCACTGCTTACAGATCTTACAGGAGAGGTATTA 483
Db 498 CGGCAATGGGCGCTTCTGTACCACTGCTTACAGATCTTACAGGAGAGGTATTA 557
QY 484 GACCTCTGGACCTGCTTGGGAGACTGCTTAATCCGAGAACTGCGGGGGAATATC 543
Db 558 GACCTCTGGACCTGCTTGGGAGACTGCTTAATCCGAGAACTGCGGGGGAATATC 617
QY 544 CTGATTCGGGTCTCTCCAGAACCCATAGTATGCTTGTGATTTTGTAGCGGCATTC 603
Db 618 CTGATTCGGGTCTCTCCAGAACCCATAGTATGCTTGTGATTTTGTAGCGGCATTC 677
QY 604 CTGCCAGCAGTCCGAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGC 663
Db 678 CTGCCAGCAGTCCGAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGC 737
QY 664 AGTCATGCTGTGCTCTGCTCAAGGTGGACCGAGCGGGAACGTTTGGGCCCCATTTGCGCAG 723
Db 738 AGTCATGCTGTGCTCTGCTCAAGGTGGACCGAGCGGGAACGTTTGGGCCCCATTTGCGCAG 797
QY 724 CGAGAGGGAACAACTCTACCTGATTGACTTGGTGGGTGAGAGCAACCGGCGCACAGGC 783
Db 798 CGAGAGGGAACAACTCTACCTGATTGACTTGGTGGGTGAGAGCAACCGGCGCACAGGC 857
QY 784 AACAGGGCCCTTGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCGGC 843
Db 858 AACAGGGCCCTTGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCGGC 917
QY 844 AAAGTGGTAGATGCGCTGAATCAGGGGCTCCCTGCTGTAACCTTATCGGGACAGCAAGCTC 903
Db 918 AAAGTGGTAGATGCGCTGAATCAGGGGCTCCCTGCTGTAACCTTATCGGGACAGCAAGCTC 977
QY 904 ACTCGCTTATTGAGGACTCTCTGGTGGTGCACGCCACAGATATCTTATGCGCAACATT 963
Db 978 ACTCGCTTATTGAGGACTCTCTGGTGGTGCACGCCACAGATATCTTATGCGCAACATT 1037
QY 964 GCGCTGAGAGACGCTTCTACCTAGACACAGCTCCGCACTCAACTTTGCTGCCAGGTCC 1023
Db 1038 GCGCTGAGAGACGCTTCTACCTAGACACAGCTCCGCACTCAACTTTGCTGCCAGGTCC 1097
QY 1024 AAGGAGGTGATCAATCG 1040
Db 1098 AAGGAGGTGATCAATTTG 1114
RESULT 10
US-09-883-096-1
; Sequence 1, Application US/09883096
; Patent No. 6680369
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Craven, Andrew
; APPLICANT: Yu, Ming
; APPLICANT: Sakowicz, Roman
; APPLICANT: Patel, Umesh A.
; APPLICANT: Davies, Katherine A.
; TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE
; FILE REFERENCE: 020552-001410US
; CURRENT APPLICATION NUMBER: US/09/883,096
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 09/594,655
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4108
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Nucleic acid sequence of human kinesin motor
; OTHER INFORMATION: protein gene Hskip3a (Figure 1).
; OTHER INFORMATION: Description of Artificial Sequence: Hskip3a gene.
US-09-883-096-1
Query Match 9.5%; Score 138.6; DB 4; Length 4108;
Best Local Similarity 51.0%; Pred. No. 4.8e-30;
Matches 448; Conservative 0; Mismatches 404; Indels 27; Gaps 4;
QY 196 TTTGATGCTTCTATGGGAGGAGTACTCAGAGGAGACATCTATGACAGTTTCAGTGCAG 255
Db 344 TTTGACCGGCTCTTGGCGAGGCGGCCACCCACAGGAGCGTGTTCAGCACACCGCAC 403
QY 256 CCCATCCTAAGGCACTTGTGGAAGGCGAGATGCGAGTGTGCTTCCCTATGAGCCACACA 315
Db 404 AGCGTCTTGAGAGCTTCTCCAGGCTACACTGCTAGTGTTCCTACGGGGCCACC 463
QY 316 GGAAGCTGGGAAGACGACCAATGCTGGGAGCCACAGAGCAACCTGGGTGATCCCGCGG 375
Db 464 GGGGCTGGGAAGACACACACCATGCTGGGAAGGGAGGGGACCCCGGCATCATGTACTG 523
```

```

Qy 376 GCTCTCATGGACCTCTCTGAGTCTCAAGGAGGAGGGTGGCGAGGGCGGCGCATGGGCC 435
Db 524 AC-----CACGTGGAACCTGTACAGGGCGCTGGAGGCGCGCCAGCAGGAGAACAC 574
Qy 436 CTTTCTGTCAACATGCTCTTACCTAGAGATCTACAGAGAGAGGTATTATAGACTCTCTGGAC 495
Db 575 TTGAGGTGCTCATCAGCTACAGGAGGTGTATATGAAACAGATCCATGACCTCTCTGGAG 634
Qy 496 CTTCTCCGAGAGCCCTGTTAATCCGAGAAGACTGCCGGGGGAATATCTCTGATTCCGGGT 555
Db 635 C---CRAAGGGCGCCCTTGGCCATCCGCGAGGACCCGACAGGGGGTGTGTGTCAGGA 691
Qy 556 CTCTCCGAGAGCCCATCAGTAGCTTGTCTGATTTTGAAGGCGACTTCTCTGCCACCCAGT 615
Db 692 CTTTCTTTCCACAGCCAGCCTCAGCCGAGCAGTGTCTGAGATATGACACAGGGGGAAC 751
Qy 616 CGAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGCGAGTCACTGCTGTG 675
Db 752 CGTAACCGCAGCAGCACCACCTGATGCCAACGGAGCTTCTCTCCGCTCCCATGCCATC 811
Qy 676 CTCCTGTGCAAGGTGGACAGCGGGAACGTTTGGCCCCCATTTCCGCCAG-----CGAGAG 729
Db 812 TTCAGATCTTTGTGAAGCAGCAGGACCGGGTTCAGGACTGACCCAGGCTGTCCAGGTG 871
Qy 730 GGAAACTCTACCTGATTGACTTGGCTGGGTGAGAGCAACACCGCGGACAGGCAACAAG 789
Db 872 GCCAAGATGAGCTTGTGACTTGGCTGGCTCAGAGCGGGCATCCAGCACCCTATGCGAAG 931
Qy 790 GGCCTTCGGCTTAAAGAGGTGAGCCATCAACACCTCCCTGTTTGTCTTGGGCAAGTG 849
Db 932 GGGGAGCGGCTGCGGGAGGGGGGCAACATCAACCGCTCTCTGCTGGCGCTCATCAACGTC 991
Qy 850 GTAGATCGCTG-----AATCAGGGCTCTCTGCTGTACCTTATCGGAGCAGCAAG 900
Db 992 CTCATGCTTGGCGGATGCAAGGGCGGCAAGCCCATGCTGCTTCCAGGAGCAGCAA 1051
Qy 901 CTACTCGCTATTGAGGACTCTCTGGGTGGCTCAGCCACAGTATCTCTTATTGCAAC 960
Db 1052 CTGACCGGCTGCTCAAGAATCTCCCTCGGGGCAACTCCGCGCAGTGTATGCTGCTGCC 1111
Qy 961 ATTGCCCTTGAGAGCGCTTACTAGACAGCTCTCCGACTCAACTTGTGCTGCCAGG 1020
Db 1112 ATAGCCCTTCCAGCTGACCTACGAGGACAGTACAAACCCCTCAAAATATGCGCGCG 1171
Qy 1021 TCCAAGAGGTGATCAATCGGCTTTTACCAATGAGAGC 1059
Db 1172 GCCAAGGATCAGGCTCTCGCTGAAGAGCAATGTGACC 1210

```

RESULT 11

```

US-09-883-096-6
; Sequence 6, Application US/09883096
; Patent No. 6680369
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Craven, Andrew
; APPLICANT: Yu, Ming
; APPLICANT: Sakowicz, Roman
; APPLICANT: Patel, Umesh A.
; APPLICANT: Davies, Katherine A.
; TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE
; FILE REFERENCE: 020552-001410US
; CURRENT APPLICATION NUMBER: US/09/883,096
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 09/594,655
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1152
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

```

```

; OTHER INFORMATION: Description of Artificial Sequence: HsKip3a
; OTHER INFORMATION: fragment
; OTHER INFORMATION: Nucleotide sequence of HsKip3a fragment used in
; OTHER INFORMATION: ATPase assay (Figure 5).
US-09-883-096-6

```

```

Query Match          9.3%; Score 136.6; DB 4; Length 1152;
Best Local Similarity 51.2%; Pred. No. 9.5e-30;
Matches 437; Conservative 0; Mismatches 389; Indels 27; Gaps 4;

```

```

Qy 196 TTTGATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGAGAGTTTCAAGTGAG 255
Db 214 TTTGACCGGGTCTTTTGGCGAGGCGGCCACCCACAGGACGTTTCCAGCACACACGCGAC 273
Qy 256 CCCATCTTAAGGCATTTGCTGGAAGGCGAGATCCAGTGTGCTTGGCTATGGACCCACACA 315
Db 274 AGCTCTCTGGACAGCTTCTCCAGGGCTACAACTGCTCAGTGTGTTGCTAGCGGGCCACC 333
Qy 316 GGAGCTGGGAGAGCGACACAAATGCTGGGCGAGCCCGAGAGCAACCTGGGGTGTATCCCGCG 375
Db 334 GGGGCTGGGAAAGACACACACCATGCTGGAAAGGAGGGGACCCCGGCATCATGTA----- 389
Qy 376 GCTCTCATGACCTCTCTGAGCTCAAGGAGAGAGGTGCGAGGGCGGCGCATGGGCC 435
Db 390 -----CCTGACCACTCGTGGAACTGTACAGGCGCTGGAGGCCCGCGCAGCAGAGAGAC 444
Qy 436 CTTTCTGTCAACCATGCTTACCTAGAGATCTTACCAGGAGAGGTATTATGACCTCTCTGAC 495
Db 445 TTGAGGTGCTCATCAGCTACCAGGAGGTGTATTAATGAACAGATCCATGACCTCTCTGGAG 504
Qy 496 CTTGCTTCCGGGAGACCTTGGTAATCCGAGAAAGACTGCGGGGGGAATATCTCTGATTCGGGT 555
Db 505 C---CCAAGGGGCGCTTGGCATCCGCGAGGACCCCGACAGGGGTGGTGGTGCNAGGA 561
Qy 556 CTCTCCAGAGGCCATCAGTAGCTTGTGCTGATTTTGAAGGCGCACTTCTGCGGCGCAGT 615
Db 562 CTTTCTTTCACAGCAGCCAGCTCAGCGAGCAGCTGCTGGAGATATCTGACCGGGGGAAC 621
Qy 616 CGAAATCGGACTGTAGGAGCCACCGGCTCAACCGAGCGCTCTCCGCGAGTCACTGCTGTG 675
Db 622 CGTAACCGCAGCAGCAACCCACTGATGTCGAAGCGGACTTCTCTCCGCTCCCATGCCATC 681
Qy 676 CTCCTGTCTAAGGTGGACAGCGGGAACGTTTGGCCCCATTTTCGCCAG-----CGAGAG 729
Db 682 TTCAGATCTTTGTGAAGCAGCAGGACCGGGTTCAGGACTGACCCAGGCTGTCCAGGTG 741
Qy 730 GGAAACTCTTACTGATGACTTGGCTGGGTGAGAGCAACCGGCGCAGGCAACAAG 789
Db 742 GCCAAGATGAGCCTGATTTGACCTGGCTGGCTCAGAGCGGGCATCCAGCACCCTATGCGAAG 801
Qy 790 GGCCTTCCGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAGTG 849
Db 802 GGGGAGCGGCTGCGGGAGGGGGGCAACATCAACCGCTCTCTGCTGGCGCTCATCAACGTC 861
Qy 850 GTAGATGCGCTGAATCAGG-----GCTCTCTCTGCTGCTTACCTTATCGGAGCAGCAAG 900
Db 862 CTCATAGCTTGGCCGATGCAAGGGCGGCAAGCCCATGTGCTTACCGGAGCAGCAA 921
Qy 901 CTCACCTGCTTATGAGGACTCTCTGGGTGGCTCAGCCCAAGATATCTTATTGCAAC 960
Db 922 CTGACCGGCTGCTCAAGAGACTCCCTCGGGGCAACTGCGGCGCAGTGTATGCTGCTGCC 981
Qy 961 ATTGCCCTTGAGAGCGCTTCTACTAGACAGTCTCCGCACTCAACTTGTGCTGCCAGG 1020
Db 982 ATCAGGCCCTTCCAGCCTTGACCTACGAGGACAGCTGACAAACCCCTCAAAATATGCGCGCG 1041
Qy 1021 TCCAAGGAGGTGA 1033
Db 1042 GCCAAGGAGATCA 1054

```

```

RESULT 12
US-09-621-233-1

```

QY	904	ACTCGCTATTGTCAGGACTCTCTGGGTGGCTCAGCGCACAGTAGTATCTCTTATTGGCAACATT	963
DB	760	ACCCGGCTCTGAAGAGACTCTCTGGGAGGAAACAGCGCACAGTGTATCGCTCACATC	819
QY	964	GCCCTCTGAGAGACGCTTCTACCTTAGACACAGTCTCCGCACTCAACTTTGTCTGCCAGTCC	1023
DB	820	AGTCTTGGAGCAGTGTCTTCCAGAGGAGTCCCGGAACACCCCTGACCTACGCCGCGCGGCCC	879
QY	1024	AAGGAGGTGATCAATCGG	1041
DB	880	ARGAACATTAAAGACTAGG	897
RESULT 13			
US-09-724-508-1			
; Sequence 1, Application US/09724508			
; Patent No. 6368841			
; GENERAL INFORMATION:			
; APPLICANT: Beraud, Christophe			
; APPLICANT: Freedman, Richard			
; TITLE OF INVENTION: No. 6368841el motor proteins and methods for			
; TITLE OF INVENTION: their use			
; FILE REFERENCE: 1045			
; CURRENT APPLICATION NUMBER: US/09/724,508			
; CURRENT FILING DATE: 2000-11-27			
; PRIOR APPLICATION NUMBER: 09/621,233			
; PRIOR FILING DATE: 2000-07-21			
; NUMBER OF SEQ ID NOS: 2			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 1			
; LENGTH: 897			
; TYPE: DNA			
; ORGANISM: Human			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (472)...(495)			
; OTHER INFORMATION: n = a, c, t, or g			
US-09-724-508-1			
Query Match 8.8%; Score 128.8; DB 3; Length 897;			
Best Local Similarity 50.1%; Pred. No. 1.6e-27;			
Matches 430; Conservative 0; Mismatches 356; Indels 72; Gaps 3			
QY	190	TACAGATTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCAGGTTCA	249
DB	106	TACCTGTTGCAGCTGGGCTTTGACTTCCGCCACCCAGGAGATGGTGTATCAGGCCACC	165
QY	250	GTGCAGCCCATCTTAAGGCACCTTGCTGGAAGGCGAGAATGCGCAGTGTCTTGGCTATGGA	309
DB	166	ACCAAGACCTATCGAGGGGCTCATCTCAGGCTACATGCCACTGTCTTTGGCTATGGC	225
QY	310	CCCA CAGAGAGTGGGAAGACGCA CACAATGTCTGGGACGCCAGAGCAACCTGGGGTGATC	369
DB	226	CCCA CAGGCTGTGGGAAACCTTACACCA TGCTGGGCA CAGACCAGGAGCCTGGCATCTAT	285
QY	370	CCGCGGCTCTCATGGACCTCCTGCAGCTCA CAGGAGGAGGGTGC CCGAGGCCGGCCA	429
DB	286	GTTCAGACCTTCAACGACCTCTTCGTGCCATCGAGGAGACCAGCAATGA	336
QY	430	TGGGCCCTTCTGTCCACCATGTCTTACCTAGAGATCTTACAGGAGAA GTATTAGACCTTC	489
DB	337	ATGGAGTATGAGTCTCCATGTCTTACCTGGAGATCTCAATAGATGATATCCGGGACCTG	396
QY	490	CTGGACCTTCTTCTGGGAGACCTGTGTAATCCGAGAGACTGCCGGGGGAA TATCTCTGATT	549
DB	397	CTGAACCCCTCCCTGGGCTA CCTGAGCTGCGGGAGGACTCTTACGGGGGTGATCCAGGTG	456
QY	550	CCGGGTCTCTCCGAAGCCCATCATAGTCTTGTGATTTTGTAGCGGCACCTTCCTGCCA	609
DB	457	GCCGCATCA CCGAANN	516
QY	610	GCCAGTCGAAATCGCATCTGTAGGAGCCACCCGGCTCAACAGCGGCTCTCCCGCAGTCAT	669

Db 517 GGGAAACGGGACAGAGACCCAGAGAGCCAGCGCGCCCAACCAAGAGTCTCTCCGCTCCCAAC 576
Qy 670 GCTGTGCTCTCTGCTCAAGGTGGACCCAGCGGGAAGTTTGGGCCCATTTGCGCCAGGAG 729
Db 577 GCGGGCGCCTGTTCA-----592
Qy 730 GGAATACTCTACCTGATTGACTTGGCTGGGTGAGAGGACCAACCGCGGACAGGCAACAAG 789
Db 593 -----TGATCGACCTGGCTGGCTCAGAGCGGCTCGCAGACACAGAAATCGT 639
Qy 790 GGCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAAAGTG 849
Db 640 GGGCAGGCTATGAAGAGGGGGCCACATCAACCGCTCACTGTGGCACTGGGCAACTGC 699
Qy 850 GTAGATCGCTGAATCAG-----GGCTCCCTCGTGTACCTTATCGGGACAGCAAGCTC 903
Db 700 ATCAACGCTTACGAGCAAGGTAGCAACAGTACATCACTATCGGACAGCAAGCTC 759
Qy 904 ACTGCGCTATTCAGGAGTCTCTGGGTGGTTCAGCCCAAGTATCTTATTGCAACATT 963
Db 760 ACCGGCTCTGAAGGACTCTCTGGAGGAAACAGCGGCAAGTATGATCGCTCACATC 819
Qy 964 GCGCTGAGAGAGCTTCTACCTAGACACAGTCTCCGACTCAACTTTGCTGCCAGTCC 1023
Db 820 AGTCTCGAGAGTGTCTTGGAGAGTCCGGAAACACCTTGACTAGCGCGCGCGGCC 879
Qy 1024 AAGGAGTGTATCAATCGG 1041
Db 880 AAGAACATTAAAGTAGG 897

RESULT 14

US-09-724-516-1
; Sequence 1, Application US/09724516
; Patent No. 6391573
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1045
; CURRENT APPLICATION NUMBER: US/09/724,516
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/621,233
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 897
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (472)...(495)
; OTHER INFORMATION: n = a, c, t, or g
US-09-724-516-1

Query Match 8.8%; Score 128.8; DB 3; Length 897;
Best Local Similarity 50.1%; Pred. No. 1.6e-27;
Matches 430; Conservative 0; Mismatches 356; Indels 72; Gaps 3;

Qy 190 TACCAGTTTCATCGCTTCTATGGGAGGAGTACTCAGCAGGACATCTATGCGAGTTCA 249
Db 106 TACTGTTTCAGCTGGCTTGTGACTTCCCGCACCCAGGAGATGGTGTATCAGGCCACC 165
Qy 250 GTGACGCCATCTTAAGGCACTTGTCTGGAAGGGCAGATCCAGTGTGCTTGTGCTATGGA 309
Db 166 ACCAAGAGCTCATCGAGGGCGTCACTCAGGCTACATGCCACTGTCTTTGCTATGGC 225
Qy 310 CCACAGGAGCTGGGAGAGCGCACAAATGCTGGGAGCCCGCAGCAACCTGGGGTGATC 369
Db 226 CCCACAGGCTGTGGGAAACCTACACATGCTGGGACAGACACAGGAGCGCTGTCATCTAT 285

Qy 370 CCGCGGCTCTCATGGACCTCTCGACGTCACRAAGGAGGAGGTGCGAGGCGCGGCCA 429
Db 286 GTTCAGACCTCTCAAGACCTCTCTCCGTGCCATCGAGGAGACCAAGATGA-----C 336
Qy 430 TGGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTACCAAGGAGAAGGTATTAGACTC 489
Db 337 ATGAGATATGAGGTCTCCATGTCTCTACCTGAGATCTACAATGAGATGATCCGGGACCTG 396
Qy 490 CTGACCCCTCTTCGGGAGACCTGGTAATCGGAGAACTCGGAGGAGGGAATATCTGTATT 549
Db 397 CTGAACCCCTCCCTGGGCTACTTGGAGCTCGGGAGGACTCTAAGGGGGTGTATCCAGGTG 456
Qy 550 CCGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGTCTGATTTTGAGCGCACTTCTCTGCCA 609
Db 457 GCGGCATCAACGANN 516
Qy 610 GCCAGTGAATCGGACTGTAGGAGCCACCGGCTCAACGAGGCTCTCTCCGCAAGTCA 669
Db 517 GGGAAACCGGACAGGAGACCCAGGAGCCCAACCGCGCCCAACAGACGCTCTCTCCGCTCCAC 576
Qy 670 GCTGTGCTCTCTGCTCAAGGTGGACCGGGAAGCTTTGGCCCCATTTGCGCAGCGAGAG 729
Db 577 GCGGGCGGCTCTGTTCA-----592
Qy 730 GGAATACTCTACCTGATTGCTTGGCTGGGTGAGAGCAACCGGCGCACAGGCAACAAG 789
Db 593 -----TGATCGACCTGGCTGGCTCAGAGCGGCTCGCAGACACAGAAATCGT 639
Qy 790 GGCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTTGTCTGGGCAAAAGTG 849
Db 640 GGGCAGCTATGAAGGAGGGGCGCCACATCAACGCTCACTGTGGCACTGGGCAACTGC 699
Qy 850 GTAGATGGCTGAATCAG-----GGCTCCCTCGTGTACCTTATCGGAGACAGCAAGCTC 903
Db 700 ATCAACGCTTACGAGCAAGGTAGCAACAGTACATCAACTATCGCAGACAGCAAGCTC 759
Qy 904 ACTCGCTATTGAGGACTCTCTGGGTGGTTCAGCCCAAGTATCTTATTGCAACATT 963
Db 760 ACCGGCTCTGAAGGACTCTCTGGAGGAAACAGCGGCAAGTATGATCGCTCACATC 819
Qy 964 GCGCTGAGAGAGCTTCTACCTAGACACAGTCTCCGACTCAACTTTGCTGCCAGTCC 1023
Db 820 AGTCTCGAGAGTGTCTTGGAGAGTCCCGGAAACACCTTGACTAGCGCGCGCGGCC 879
Qy 1024 AAGGAGTGTATCAATCGG 1041
Db 880 AAGAACATTAAAGTAGG 897

RESULT 15

US-10-090-695-1
; Sequence 1, Application US/10090695
; Patent No. 6664072
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1045
; CURRENT APPLICATION NUMBER: US/10/090,695
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: 09/621,233
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 897
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (472)...(495)

; OTHER INFORMATION: n = a, c, t, or g
US-10-090-695-1

```
Query Match      8.8%; Score 128.8; DB 4; Length 897;
Best Local Similarity 50.1%; Pred. No. 1.6e-27;
Matches 430; Conservative 0; Mismatches 356; Indels 72; Gaps 3;

QY 190 TACAGTTTGTGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCAGTTCA 249
DB |||||
QY 106 TACCTGTTGAGCTGTCCTTTCACCTCCACCCAGGAGATGGTGTATCAGGCCACC 165
DB |||||
QY 250 GTGCAGCCCATCTTAAGGCACCTTGTGGAAGGCGAGATGCCAGTGTCTTGCCTATGGA 309
DB |||||
QY 166 ACCAAGAGCCTCATCAGGGGCGTCATCTCAGGCTACAATGCCACCTGTCTTTGCTATGSC 225
DB |||||
QY 310 CCCACAGGAGCTGGGAGAGCGACACAATGCTGGGCGAGCCACAGAGCAACCTGGGGTGATC 369
DB |||||
QY 226 CCCACAGGCTGTGGGAGAACTTACACCATGCTGGGCACAGACCAGGAGCCTGGCATCTAT 285
DB |||||
QY 370 CCGCGGGCTCTCATGAGACCTCTCTGACGCTCAAGGGAGGAGGGTCCGAGGGCCGGCCA 429
DB |||||
QY 286 GTTCAGACCTCAACGACCTCTTCCTGCGCATCGAGGAGACCAGCAATGA-----C 336
DB |||||
QY 430 TGGGCCCTTCTGTCTACCATGCTTACTTAGAGATCTACCAGGAGAGGTATTAGACCTC 489
DB |||||
QY 337 ATGGAGTATGAGGTCTCCATGTCCTACCTGGAGATCTCAATGAGATGATCCGGGACCTG 396
DB |||||
QY 490 CTGGACCTCTGCTTCGGGAGACCTGGTAAATCCGAGAGAGACTGCCGGGGAAATATCCTGATT 549
DB |||||
QY 397 CTGAAACCCCTCCCTGGGCTACCTGGAGCTCCGGGAGGACTCTAAGGGGGTGTCCAGGTG 456
DB |||||
QY 550 CCGGGTCTCTCCAGAGCCCATCATGATGTTTGTCTGATTTTGTAGCGGCACTTCTGGCCA 609
DB |||||
QY 457 GCGGGCATCAACGAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 516
DB |||||
QY 610 GCCAGTCTGAATCGGAGCTGTAGGAGCCACCCGGCTCAACAGCGCTCTCCCGCAGTCAAT 669
DB |||||
QY 517 GGGAAACCGGAGAGAGACCCAGGAGCCCAAGGGCGCCCAACGAGACGTCCTCCGCTCCAC 576
DB |||||
QY 670 GCTGTGCTCTCTGCTCAAGTGTGAGACGAGCGGGAACGTTTGGCCCCCATTTCCGCCAGGAG 729
DB |||||
QY 577 GCGGGCGCTGTCTCA----- 592
DB |||||
QY 730 GGAAGACTCTACCTGATTGCTTGGCTGGGTGAGAGCAACCGGCGCACAGGCAACAAG 789
DB |||||
QY 593 -----TGATCGACCTGGCTGAGCGCGCTCGAGAGACAGAAATCGT 639
DB |||||
QY 790 GGCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAGTG 849
DB |||||
QY 640 GGGAGCGTATGAGAGGGGGGCCACATCAACCGCTCACTGTGGCACTGGGCAACTGC 699
DB |||||
QY 850 GTAGATCGCTGAATCAG-----GGCCTCCCTGTTGTTACTTTATCGGGACAGCAAGCTC 903
DB |||||
QY 700 ATCAAGCCCTGAGCGCAAGGGTAGCAACAAGTACATCAACTATCGGACAGCAAGCTC 759
DB |||||
QY 904 ACTGCTATGAGGAGCTCTGGGTGGCTCAGCCCAAGTATCTTATTGCCAATTT 963
DB |||||
QY 760 ACCCGGCTCTGAGGAGCTCTGAGGAGGAAACAGCGGCAAGTATGATCGCTCACATC 819
DB |||||
QY 964 GCCCTCAGAGACCTCTTACTTAGACAGTCTCCGCACTCAACTTTGCTGCCAGGTCC 1023
DB |||||
QY 820 AGTCTGCGAGCAGTGTCTGAGGAGTCCGGGAACACCTGACCTACGCGCGCGGCC 879
DB |||||
QY 1024 AAGAGGTGATCAATCGG 1041
DB |||||
QY 880 AAGAACATTAAGACTAGG 897
DB |||||
```

Search completed: November 10, 2004, 18:26:07
Job time : 137.833 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 9, 2004, 19:47:28 ; Search time 6469.64 Seconds
(without alignments)
10701.074 Million cell updates/sec

Title: US-10-797-893-7
Perfect score: 1464
Sequence: 1 atgggtcgctgcggctaag.....accattgtcccaaatgta 1464

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1464	100.0	1464	6 AR210055	AR210055 Sequence
2	1461	99.8	1538	6 AR210053	AR210053 Sequence
3	1458	99.6	1998	9 BT007259	BT007259 Homo sapi
4	1458	99.6	1998	12 BT007888	BT007888 Synthetic
5	1458	99.6	2097	6 AR304057	AR304057 Sequence
6	1458	99.6	2097	9 AB017430	AB017430 Homo sapi
7	1458	99.6	2117	9 BC028155	BC028155 Homo sapi
8	1458	99.6	2134	9 BC004352	BC004352 Homo sapi
9	1067.8	72.9	2086	10 BC003427	BC003427 Mus muscu
10	1038.4	70.9	1041	6 AR210054	AR210054 Sequence
11	1035.4	70.7	1115	6 AR210052	AR210052 Sequence
12	881.8	60.2	151041	2 AC101919	AC101919 Mus muscu
13	881.8	60.2	164759	10 AC101752	AC101752 Mus muscu
14	740.6	50.6	297639	2 AC101908	AC101908 Mus muscu
15	706.8	48.3	196674	10 AC102127	AC102127 Mus muscu
16	511.6	34.9	2756	5 BC063896	BC063896 Xenopus t
17	510.2	34.8	2583	5 XLA249841	XLA249841 Xenopus l
18	508.6	34.7	2565	5 BC043733	BC043733 Xenopus l
19	508.6	34.7	2640	5 BC073177	BC073177 Xenopus l

20	501	34.2	2593	5	AF267850	AF267850 Xenopus l
21	501	34.2	2661	5	XLA249840	XLA249840 Xenopus l
22	501	34.2	2781	5	BC070549	BC070549 Xenopus l
23	501	34.2	2819	5	AF267849	AF267849 Xenopus l
c 24	395	27.0	68916	2	AC101927	AC101927 Mus muscu
25	376.4	25.7	386	6	AX247007	AX247007 Sequence
c 26	365	24.9	174007	9	AC023831	AC023831 Homo sapi
c 27	365	24.9	174363	9	AC009133	AC009133 Homo sapi
28	361.8	24.7	3396	9	AB017333S2	AB017333 Homo sapi
29	353	24.1	105485	9	HUAC002301	AC002301 Homo sapi
30	346.4	23.7	444	10	AF013119	AF013119 Mus muscu
31	265	18.1	228783	2	AC094377	AC094377 Rattus no
c 32	265	18.1	235806	2	AC133764	AC133764 Rattus no
c 33	265	18.1	289504	2	AC123480	AC123480 Rattus no
c 34	262.4	17.9	195987	10	AC122537	AC122537 Mus muscu
c 35	262.4	17.9	226601	10	AC122863	AC122863 Mus muscu
36	208.8	14.3	1982	3	AK114179	AK114179 Ciona int
37	185.6	12.7	2095	6	AX833125	AX833125 Sequence
38	185.6	12.7	2095	9	AK094619	AK094619 Homo sapi
39	165	11.3	146415	5	AL929469	AL929469 Zebrafish
40	165	11.3	176906	2	CR354587	CR354587 Danio rer
41	153.8	10.5	3358	10	BC057614	BC057614 Mus muscu
42	145.8	10.0	435	5	MSU34658	U34658 Morone saxa
43	145.2	9.9	44004	2	AC092563	AC092563 Hyllobates
44	143	9.8	183586	9	AC092562	AC092562 Papio ham
45	139	9.5	9701	9	AB017333S3	AB017335 Homo sapi

ALIGNMENTS

RESULT 1
AR210055
LOCUS AR210055 1464 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 7 from patent US 6387644.
ACCESSION AR210055
VERSION AR210055.1 GI:21512186
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1464)
AUTHORS Beraud,C.
TITLE Motor proteins and methods for their use
JOURNAL Patent: US 6387644-A 7 14-MAY-2002;
FEATURES
source Location/Qualifiers
1..1464
/organism="unknown"
/mol_type="unassigned DNA"

Query Match	100.0%	Score 1464;	DB 6;	Length 1464;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1464;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGGGTGCTGTCGGCTAAGCAAGATTGGAGTACTCGTCGTCACCTCCAGCTCGCGTA	60	
Db	1	ATGGGTGCTGTCGGCTAAGCAAGATTGGAGTACTCGTCGTCACCTCCAGCTCGCGTA	60	
Qy	61	AGGGTGGCTGTGCGACTCGGCCATTGTGGATGGAAACAGCGGAGCAAGTGATCCCCCC	120	
Db	61	AGGGTGGCTGTGCGACTCGGCCATTGTGGATGGAAACAGCGGAGCAAGTGATCCCCCC	120	
Qy	121	TGTGTGGGGGCATGACACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACACACAGAG	180	
Db	121	TGTGTGGGGGCATGACACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACACACAGAG	180	
Qy	181	ACTCTCAATACACAGTTTGATGCTCTTATGGGGAGGAGTACTACAGAGGACATCTAT	240	
Db	181	ACTCTCAATACACAGTTTGATGCTCTTATGGGGAGGAGTACTACAGAGGACATCTAT	240	
Qy	241	GCAGGTTTCACTGACGAGCCCATCTTCTGGAAGGCGAGATGCCAGTGTGCTT	300	

Db 241 GCAGGTTACGTGAGCGCCATCTTAAGGCACTTGCTGGAGGCGCAATGCCAGTGTGCTT 300
Qy 301 GCTATGAGCCACAGGAGCTGGGAGAGCGCACAAATGCTGGGAGGCCCGCAGCAACCT 360
Db 301 GCTATGAGCCACAGGAGCTGGGAGAGCGCACAAATGCTGGGAGGCCCGCAGCAACCT 360
Qy 361 GGGGTGATCCGGGGCTCTATGAGCTCTCTGAGCTCTCTGAGCTCAAGGAGGAGGCTCCGAG 420
Db 361 GGGGTGATCCGGGGCTCTATGAGCTCTCTGAGCTCTCTGAGCTCAAGGAGGAGGCTCCGAG 420
Qy 421 GCGCGCCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTTACAGAGAGGTA 480
Db 421 GCGCGCCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTTACAGAGAGGTA 480
Qy 481 TTAGACCTCTGAGCCCTGCTTGGGAGACCTGTAATCCGAGAGAGCTGCGGGGGAAT 540
Db 481 TTAGACCTCTGAGCCCTGCTTGGGAGACCTGTAATCCGAGAGAGCTGCGGGGGAAT 540
Qy 541 ATCTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCAC 600
Db 541 ATCTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCAC 600
Qy 601 TTCCTCCAGCAGTCGAAATCGACTGTAGAGCCACCGGCTCAACAGCGCTCTCC 660
Db 601 TTCCTCCAGCAGTCGAAATCGACTGTAGAGCCACCGGCTCAACAGCGCTCTCC 660
Qy 661 CGCAGTCATGCTGCTCTGCTCAAGGTGAGCCAGCGGGAACGTTTGGCCCATTTGCG 720
Db 661 CGCAGTCATGCTGCTCTGCTCAAGGTGAGCCAGCGGGAACGTTTGGCCCATTTGCG 720
Qy 721 CAGCGAGAGGAAATCTTACCTGATTTGCTGGCTGAGGAGCAACCGGCGACA 780
Db 721 CAGCGAGAGGAAATCTTACCTGATTTGCTGGCTGAGGAGCAACCGGCGACA 780
Qy 781 GGCACAAAGGCGCTTCGGCTAAAGAGAGTGAGCCATCAACACCTCCCTGTTTGTCTG 840
Db 781 GGCACAAAGGCGCTTCGGCTAAAGAGAGTGAGCCATCAACACCTCCCTGTTTGTCTG 840
Qy 841 GGCACAAAGTGTAGATGCGCTGAATCAAGGCTCTCTGCTGCTGCTTATCGGAGCAAG 900
Db 841 GGCACAAAGTGTAGATGCGCTGAATCAAGGCTCTCTGCTGCTGCTTATCGGAGCAAG 900
Qy 901 CTCACTCGCTTATGAGGAGCTCTCTGGTGGCTCAGCCACAGTATCTTATGCGAAC 960
Db 901 CTCACTCGCTTATGAGGAGCTCTCTGGTGGCTCAGCCACAGTATCTTATGCGAAC 960
Qy 961 ATTGCCCCGAGAGAGCTTTCTACCTAGACACAGTCTCGACATCAACTTTGCTGCGAG 1020
Db 961 ATTGCCCCGAGAGAGCTTTCTACCTAGACACAGTCTCGACATCAACTTTGCTGCGAG 1020
Qy 1021 TCCAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTGAGGCTCATGCTTGGGA 1080
Db 1021 TCCAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTGAGGCTCATGCTTGGGA 1080
Qy 1081 CTTGTTAAGCTGCTCAGAAAGAAATTCCTTGGTCCACAGAGGCAAGAGCCGAGGC 1140
Db 1081 CTTGTTAAGCTGCTCAGAAAGAAATTCCTTGGTCCACAGAGGCAAGAGCCGAGGC 1140
Qy 1141 CTTGAGGAAGAGGAGATTTGGAGCCCTGAGCCATGAGAGCTCAGGCTCTGCTCCAG 1200
Db 1141 CTTGAGGAAGAGGAGATTTGGAGCCCTGAGCCATGAGAGCTCAGGCTCTGCTCCAG 1200
Qy 1201 AAATCAGCCCTTACAGAGCTTAAGCAGATGAGCCCGGCTGCTGAGGAGCCCTCCTC 1260
Db 1201 AAATCAGCCCTTACAGAGCTTAAGCAGATGAGCCCGGCTGCTGAGGAGCCCTCCTC 1260
Qy 1261 AGCTTGAGCCCTGCTGCTTGGCTCCAGGAGGAGCCAGGGGCGCTCTGTTGAGTACCCCA 1320
Db 1261 AGCTTGAGCCCTGCTGCTTGGCTCCAGGAGGAGCCAGGGGCGCTCTGTTGAGTACCCCA 1320
Qy 1321 AAGCGAGAGCGGATGCTGCTTAATGAAGACAGTAGAAGAGGAGCTTAGAGATTGAGG 1380
Db 1321 AAGCGAGAGCGGATGCTGCTTAATGAAGACAGTAGAAGAGGAGCTTAGAGATTGAGG 1380

Qy 1381 CTTAAGACGAGCAAAAGAACTGGAGGCCAAGATGTTGCCAGAAAGGCTGAGGAAAAG 1440
Db 1381 CTTAAGACGAGCAAAAGAACTGGAGGCCAAGATGTTGCCAGAAAGGCTGAGGAAAAG 1440
Qy 1441 GAGAACCAATTGCTCCCACAATGTGA 1464
Db 1441 GAGAACCAATTGCTCCCACAATGTGA 1464
RESULT 2
AR210053
LOCUS AR210053 1538 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 3 from patent US 6387644.
ACCESSION AR210053
VERSION AR210053.1 GI:21512183
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1538)
AUTHORS Beraud, C.
TITLE Motor proteins and methods for their use
JOURNAL Patent: US 6387644-A 3 14-MAY-2002;
FEATURES
source 1. 1538
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 99.8%; Score 1461; DB 6; Length 1538;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 GGTCTCTGTCGGCTAAGCAAGATTGGAGCTACTGCTCTCCACTCTCCAGCTCGCTAAGG 63
Db 78 GGTCTCTGTCGGCTAAGCAAGATTGGAGCTACTGCTCTCCACTCTCCAGCTCGCTAAGG 137
Qy 64 GTGGCTGTGCACTCGCGCCATTTGTGATGGAAACAGCGGAGCAAGTGTATCCCTGT 123
Db 138 GTGGCTGTGCACTCGCGCCATTTGTGATGGAAACAGCGGAGCAAGTGTATCCCTGT 197
Qy 124 GTGCGGGCATGGAAGAGCTGCTCTAGAGATTGTAACTGGAGGAAACACAGGAGACT 183
Db 198 GTGCGGGCATGGAAGAGCTGCTCTAGAGATTGTAACTGGAGGAAACACAGGAGACT 257
Qy 184 CTCAAATACCACTTTGATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTATGCA 243
Db 258 CTCAAATACCACTTTGATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTATGCA 317
Qy 244 GCTTCAGTGCAGCCCATCTTAAGGCACTTGTCTGGAAGGCGAGATGCCAGTGTGCTGCC 303
Db 318 GCTTCAGTGCAGCCCATCTTAAGGCACTTGTCTGGAAGGCGAGATGCCAGTGTGCTGCC 377
Qy 304 TATGAGCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGCCAGAGCAACTGGG 363
Db 378 TATGAGCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGCCAGAGCAACTGGG 437
Qy 364 GTGATCCCGGGCTCTCATGACCTCTCTGAGCTCACAAGGAGGAGGCTCCGAGGCG 423
Db 438 GTGATCCCGGGCTCTCATGACCTCTCTGAGCTCACAAGGAGGAGGCTCCGAGGCG 497
Qy 424 CGGCCATGGGCGCTTCTGTCACTGCTTCTACCTAGAGATCTA CAGAGAGGATTTA 483
Db 498 CGGCCATGGGCGCTTCTGTCACTGCTTCTACCTAGAGATCTA CAGAGAGGATTTA 557
Qy 484 GACCTCTGAGCCCTGCTTGGGAGACCTGGTAATCCGAGAGACTGCCGGGGGATATC 543
Db 558 GACCTCTGAGCCCTGCTTGGGAGACCTGGTAATCCGAGAGACTGCCGGGGGATATC 617
Qy 544 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTTCAGGGGCACTTC 603
Db 618 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTTCAGGGGCACTTC 677

Qy	604	CTGCCAGCAGTCGAAATCGAATGCTAGAGCCACCCGGCTCAACACGCGCTCTCTCCGCG	663
Db	678	CTGCCAGCAGTCGAAATCGAATGCTAGAGCCACCCGGCTCAACACGCGCTCTCTCCGCG	737
Qy	664	AGTCATGCTGCTCTGCTCAAGTGTGACGAGCGGAACTTTGGCCCAATTCGCGAG	723
Db	738	AGTCATGCTGCTCTGCTCAAGTGTGACGAGCGGAACTTTGGCCCAATTCGCGAG	797
Qy	724	CGAGAGGAAACTCTACTGATGACTTGGCTGAGAGGCAACCGCGCACAGGC	783
Db	798	CGAGAGGAAACTCTACTGATGACTTGGCTGAGAGGCAACCGCGCACAGGC	857
Qy	784	AACAGGCGCTTCGGCTAAAGAGAGAGTGAGCCATCAACACCTCTCTCTGCTGGC	843
Db	858	AACAGGCGCTTCGGCTAAAGAGAGAGTGAGCCATCAACACCTCTCTCTGCTGGC	917
Qy	844	AAAGTGTAGATGCTGCTCAATCAGGCGCTCTCTGCTGACCTTATCGGACAGCAAGTC	903
Db	918	AAAGTGTAGATGCTGCTCAATCAGGCGCTCTCTGCTGACCTTATCGGACAGCAAGTC	977
Qy	904	ACTCGCTATTGAGGAGCTCTGCTGCTGCTCAGCCACAGTATCTTATTCACCAAT	963
Db	978	ACTCGCTATTGAGGAGCTCTGCTGCTGCTCAGCCACAGTATCTTATTCACCAAT	1037
Qy	964	GCCCTCGAGAGACGCTTCTACTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGTCC	1023
Db	1038	GCCCTCGAGAGACGCTTCTACTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGTCC	1097
Qy	1024	AAGAGGTGATCAATCGGCTTTTACCAATGAGAGCTGAGCCTCATGCTTGGGACCT	1083
Db	1098	AAGAGGTGATCAATCGGCTTTTACCAATGAGAGCTGAGCCTCATGCTTGGGACCT	1157
Qy	1084	GTTAAGCTGTCTCAGAAAGAAATTCCTGCTCCACAGAGGCAAGAGAGCCGAGGCT	1143
Db	1158	GTTAAGCTGTCTCAGAAAGAAATTCCTGCTCCACAGAGGCAAGAGAGCCGAGGCT	1217
Qy	1144	GAGGAAGAGGAGATTTGGAGCGCTTGAGCCATGAGCCTCAGCCTCTGCTCCAGAAA	1203
Db	1218	GAGGAAGAGGAGATTTGGAGCGCTTGAGCCATGAGCCTCAGCCTCTGCTCCAGAAA	1277
Qy	1204	CTGAGCCCTCAGAGAGCTAAGCAGATGAGACCGGCGCATGCTGGAGCGCTCTCAGC	1263
Db	1278	CTGAGCCCTCAGAGAGCTAAGCAGATGAGACCGGCGCATGCTGGAGCGCTCTCAGC	1337
Qy	1264	TTGACCGCTGCTGCTTCCAGGGAGCCAGGGGCGCTCTGTTGAGTACCCCAAAG	1323
Db	1338	TTGACCGCTGCTGCTTCCAGGGAGCCAGGGGCGCTCTGTTGAGTACCCCAAAG	1397
Qy	1324	CGAGAGCGGATGCTGCTAATGAAGACAGTGAAGAGAGGACCTTAGAGATTGAGAGGCTT	1383
Db	1398	CGAGAGCGGATGCTGCTAATGAAGACAGTGAAGAGAGGACCTTAGAGATTGAGAGGCTT	1457
Qy	1384	AAGACGAGCAAAAGAACTGGAGGCGCAAGATTTGGCCAGAGGCTGAGGAAAGGAG	1443
Db	1458	AAGACGAGCAAAAGAACTGGAGGCGCAAGATTTGGCCAGAGGCTGAGGAAAGGAG	1517
Qy	1444	AACCATTTCCCAATGTGA	1464
Db	1518	AACCATTTCCCAATGTGA	1538
RESULT 3	BT007259		
LOCUS	BT007259	1998 bp	linear
DEFINITION	Homo sapiens kinesin-like 4 mRNA, complete cds.		PRI 13-MAY-2003
ACCESSION	BT007259		
VERSION	BT007259.1		
KEYWORDS	FLI CDNA.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	AUTHORS		
TITLE			
JOURNAL			
REFERENCE	AUTHORS		
TITLE			
JOURNAL			
COMMENT			
FEATURES	source		
	1. 1998		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="GH00245X1.0"		
	/clone_lib="BD Creator (TM) CDS Library derived from MGC collection"		
	/lab_host="DH5alpha T1 resistant"		
	/notes="vector: pDNR-Dual"		
	1. 1998		
	/codon_start=1		
	/product="kinesin-like 4"		
	/protein_id="AAP35923.1"		
	/db_xref="GI:30583357"		
	/translation="MAAGSTQQRREMAAASAAAISGAGCRSLKIGATRRPPARV		
	RVARLPFVDGTAGADPPCVRMGDSCLSIANRNHQTLYQFOAFYGERSTQDD		
	IYAGSQVPLIRLHLEGONASVLAAYGTAGKTHMLGSPQPGVIFPALMDLLQUTRE		
	ADGERFPLASRNTVTGATRLNRSRSHAVLVKVDQERELAPFRQEGKLYLIDIA		
	GSNDERTNGKGLKXESGINTSLFVLGVDAALQGLPRVPYRDSKLTLLQDSLG		
	LSGSHSLIANIAPERRFYLDTSALNPAASAKVINRPFNTESLQHPALGPKVKSKE		
	LLGPPPKARAGPEEBEIGSPPEMPAASAKLSPLQKLSMDPAMLEKLLSLDRLL		
	ASQSGCAPLLSTPKRPMVLMKTVVEKDLIEIRLTKQKEAKMLAKAEKENHC		
	PTMLRPLSHRTVTGAKPLKAVVMPQLIOEQAASPNABEIHILKNKGRKKEISLDA		
	EPEKABDCWELQISPELLAHGRQKILDLNLSGARDLRSIORIGPKKQLIVGWREL		
	HGPFQVEDLERVEGITGKQMESFLKXNLIUGLAAGRCGAS"		
ORIGIN			
Query Match	99.6%;	Score 1458;	DB 9; Length 1998;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 1458;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	4	GTGCGCTGTCGGCTAAGCAAGATTGGAGCTACTGCTGCTCCACCTCCAGCTCGCGTAAG	63
Db	76	GTCGCTGTCGGCTAAGCAAGATTGGAGCTACTGCTGCTCCACCTCCAGCTCGCGTAAG	135
Qy	64	GTGCGCTGTCGGCTGCGGCCATTTGTGGATGGAAACAGCGGAGCAAGTGTATCCCTCTGT	123
Db	136	GTGCGCTGTCGGCTGCGGCCATTTGTGGATGGAAACAGCGGAGCAAGTGTATCCCTCTGT	195
Qy	124	GTGCGGCGCATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACT	183
Db	196	GTGCGGCGCATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACT	255
Qy	184	CTCAATACAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	243
Db	256	CTCAATACAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	315

244 GATTTCAGTTCAGCCATCTAAGGCATTTGCTGGAAGGCGCAATGCCAGTGTGCTTCC 303
Db
316 GATTTCAGTTCAGCCATCTAAGGCATTTGCTGGAAGGCGCAATGCCAGTGTGCTTCC 375
Qy
304 TATGGAACACAGGAGCTGGGAAGAGCACAACAATGCTGGGAGCCAGAGCAACTGGG 363
Db
376 TATGGAACACAGGAGCTGGGAAGAGCACAACAATGCTGGGAGCCAGAGCAACTGGG 435
Qy
364 GTGATCCCGGGGCTCTATGACCTCTGACCTCCAGAGCTCACAAGGAGGAGGGTCCAGGCG 423
Db
436 GTGATCCCGGGGCTCTATGACCTCTGACCTCCAGAGCTCACAAGGAGGAGGGTCCAGGCG 495
Qy
424 CGGCAATGGGCGCTTCTGTCCACCATGTCTTACCTAGAGATCTACAGGAGAGGTATTA 483
Db
496 CGGCAATGGGCGCTTCTGTCCACCATGTCTTACCTAGAGATCTACAGGAGAGGTATTA 555
Qy
484 GACCTCTGGAACCTGCTGGGAGACCTGTAATCCGAGAGACTGCGGGGGGATATC 543
Db
556 GACCTCTGGAACCTGCTGGGAGACCTGTAATCCGAGAGACTGCGGGGGGATATC 615
Qy
544 CTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTTCAGCGGCATTC 603
Db
616 CTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTTCAGCGGCATTC 675
Qy
604 CTGCCAGCCAGTCAAAATCGACTGTAGGAGCACCAGGCTCAACAGGCTCTCTCCGCG 663
Db
676 CTGCCAGCCAGTCAAAATCGACTGTAGGAGCACCAGGCTCAACAGGCTCTCTCCGCG 735
Qy
664 AGTCATGCTGTCTCTGCTCAAGTGGAGCCAGGAGCTTTGGCCCCCATTTGCCAG 723
Db
736 AGTCATGCTGTCTCTGCTCAAGTGGAGCCAGGAGCTTTGGCCCCCATTTGCCAG 795
Qy
724 CGAGAGGAAAATCTTACCTGATTGACTTGGCTGGGTCAAGGAGCAACCGGCGCAGAGC 783
Db
796 CGAGAGGAAAATCTTACCTGATTGACTTGGCTGGGTCAAGGAGCAACCGGCGCAGAGC 855
Qy
784 AACAGGGCTCTGGCTTAAAGAGAGTGGAGCATTCAACCTCCCTGTTGCTTGGGCG 843
Db
856 AACAGGGCTCTGGCTTAAAGAGAGTGGAGCATTCAACCTCCCTGTTGCTTGGGCG 915
Qy
844 AAGTGTAGATGCTGCTGAATCAGGCTCCCTGCTGATCTTATCGGACACCAAGCTC 903
Db
916 AAGTGTAGATGCTGCTGAATCAGGCTCCCTGCTGATCTTATCGGACACCAAGCTC 975
Qy
904 ACTCGCTTATGACGACTCTCTGGGTGGTCAAGCCAGCAGTATCTTATGTCACATTC 963
Db
976 ACTCGCTTATGACGACTCTCTGGGTGGTCAAGCCAGCAGTATCTTATGTCACATTC 1035
Qy
964 GCGCTTGAGAGACGCTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGTCC 1023
Db
1036 GCGCTTGAGAGACGCTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGTCC 1095
Qy
1024 AAGAGGTGATCAATCGGCTTTTACCAATGAGAGCCTGACGCTCATGCTTGGGACCT 1083
Db
1096 AAGAGGTGATCAATCGGCTTTTACCAATGAGAGCCTGACGCTCATGCTTGGGACCT 1155
Qy
1084 GTTAAGTGTCTCAAGAAATGCTTGGTCTCAGAGGCAAGAGAGCCCGGCGCT 1143
Db
1156 GTTAAGTGTCTCAAGAAATGCTTGGTCTCAGAGGCAAGAGAGCCCGGCGCT 1215
Qy
1144 GAGGAGGAGATGGGAGCCCTGAGCCCATGCTCCAGCTCTGCTCCAGAAA 1203
Db
1216 GAGGAGGAGATGGGAGCCCTGAGCCCATGCTCCAGCTCTGCTCCAGAAA 1275
Qy
1204 CTCAGCCCTCAAGAGCTAAGCAGCATGAGCCCGGCTATGCTGGAGCGCTCTCAGC 1263
Db
1276 CTCAGCCCTCAAGAGCTAAGCAGCATGAGCCCGGCTATGCTGGAGCGCTCTCAGC 1335
Qy
1264 TTGGAACGCTGCTGCTCCAGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1323
Db
1336 TTGGAACGCTGCTGCTCCAGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1395
Qy
1324 CGAGAGCGGATGGTCTTAATGAAGAAGTGAAGAGAGGAGGACCTAGAGATTGAGAGGCTT 1383

1396 CGAGAGCGGATGGTCTTAATGAAGAAGTGAAGAGAGGACCTAGAGATTGAGAGCTT 1455
Qy
1384 AAGACGAAGCAAAAGAACTGGAGGCCAAGATGTGGCCAGAGGCTGAGGAAAAGAG 1443
Db
1456 AAGACGAAGCAAAAGAACTGGAGGCCAAGATGTGGCCAGAGGCTGAGGAAAAGAG 1515
Qy
1444 AACATTGTCCCAATG 1461
Db
1516 AACATTGTCCCAATG 1533

RESULT 4

BT007888

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1998 bp mRNA linear SYN 13-MAY-2003
Synthetic construct Homo sapiens kinesin-like 4 mRNA, partial cds.BT007888
BT007888
BT007888.1 GI:30584614
FLI CDNA.synthetic construct
synthetic construct
artificial sequences.1 (bases 1 to 1998)
Kandinya, M., Chen, X., Rolf, A., Halleck, A., Hines, L., Eisenstein, S.,
Kandinya, M., Raphael, J., Moreira, D., Kelley, T., LaBaer, J., Lin, Y.,
Phelan, M. and Farmer, A.Cloning of human full-length CDSs in BD Creator(TM) System Donor
vectorUnpublished
2 (bases 1 to 1998)
Kandinya, M., Chen, X., Rolf, A., Halleck, A., Hines, L., Eisenstein, S.,
Kandinya, M., Raphael, J., Moreira, D., Kelley, T., LaBaer, J., Lin, Y.,
Phelan, M. and Farmer, A.Direct Submission
Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow
Circle, Palo Alto, CA 94303, USAThis CDS clone is a part of a collection of human full length
expression clones generated by BD Biosciences Clontech and the
Harvard Institute of Proteomics. Each CDS has been cloned in two
forms: with and without stop-codon (to allow fusion with C-terminal
tag). The CDS has been directionally cloned using BD In-Fusion(TM)
cloning system between the SalI and HindIII sites of the pDNR-DUAL
vector. Additional sequences in the clone: 'ACC' after SalI site
and before 'ATG' to provide Kozak consensus sequence; 'GG' after
last codon and before HindIII site to maintain reading frame.

Clone distribution: http://bioinfo.clontech.com/orfclones.

Location/Qualifiers
1. .1998
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/clone="CH0024511.0"
/clone_lib="BD Creator(TM) CDS Library derived from MGC
collection"
/lab_host="DH5alpha Tl resistant"
/note="Vector: pDNR-Dual"1. .>1998
/note="Mutations: 1997:Stop->Leu"
/codon_start=1
/transl_table=11
/product="Homo sapiens kinesin-like 4"
/protein_id="AAP36560.1"
/db_xref="GI:30584615"/translation="MAAGSTQQRREMAAASAAATSGAGRCRLSKIGATRRPPARV
RVARLRPFVDGTAGADPPCVGRMDSCLSEIANWNRHETLYQFDAYGSRSTOOD
IYAGSVQPIRLHLEQNAVLAYGPTGAKTHLMGSPQGVTPRALMDLLQLTRE
EGAGRPWALSVTMSLEYIQKYLDDLPASGDLVREDCRGNLI1PGLSOKPISFF
ADPREFLPASRNRTVGTARNQSRSHAVLLVKVDQERLAPRQRGKLYLIDLA
GSDNRTGNKGLRKESGINTSLFVLGVKVDALNQGLRPVPRPSKRLRLQDSIG
GSAHSLIANIAPRRFYDTSALNFAARSKVINRPTNSLOPHALPGLVKLSQKE
LUGPPKARGPPEEIEIPRMAAPASQKLSPLQKLSMDPAMRLSLDRLL
ASGSGAPLSTPKERVMKTVBEKLETERLTKQKLEAKMLAKAEKNEHC
PTWLRPLSHRTVTGAKPLKAVVMPQLIQEQAAASPNABIIHLKNNKRRKRLSLDAL

FEATURES

source

CDS

EPBEKAEBCWELQISPELLAHGROKILDLNLEGSARDLRSIQRIPKKAQLIVGREL
HGPFQVEDLSEVBTGKQMESFLKILGLAAGRCASL"

ORIGIN

Query Match 99.6%; Score 1458; DB 12; Length 1998;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 4 GGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGGTAAGG 63
DB 76 GGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGGTAAGG 135
QY 64 GTGGCTGTGCACTCGGCCCAATTTGTGATGGAACAGCGGAGCAAGTATCCCCCTGT 123
DB 136 GTGGCTGTGCACTCGGCCCAATTTGTGATGGAACAGCGGAGCAAGTATCCCCCTGT 195
QY 124 GTGGGGGCATGGACAGCTGCTCTAGAGATTGCTAACTGGAGAACCAACAGAGACT 183
DB 196 GTGGGGGCATGGACAGCTGCTCTAGAGATTGCTAACTGGAGAACCAACAGAGACT 255
QY 184 CTCAAAATACCAATTTGATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 243
DB 256 CTCAAAATACCAATTTGATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 315
QY 244 GGTTCAGTGCAGCCCATCTTAAGGCACTTGTCTGGAAGGGCAGAAATGCCAGTGTGCTGCC 303
DB 316 GGTTCAGTGCAGCCCATCTTAAGGCACTTGTCTGGAAGGGCAGAAATGCCAGTGTGCTGCC 375
QY 304 TATGACCCACAGGAGCTGGGAAGCAGCACCAATGCTGGGAGAGGGGTCCGAGGGC 363
DB 376 TATGACCCACAGGAGCTGGGAAGCAGCACCAATGCTGGGAGAGGGGTCCGAGGGC 435
QY 364 GTGATCCCGGGCTCTCATCGAGCTCTCTGAGCTTCAAGAGGAGGGGTCCGAGGGC 423
DB 436 GTGATCCCGGGCTCTCATCGAGCTCTCTGAGCTTCAAGAGGAGGGGTCCGAGGGC 495
QY 424 CGGCATGCGGGCTTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGAAAGTATTA 483
DB 496 CGGCATGCGGGCTTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGAAAGTATTA 555
QY 484 GACCTCTGGACCTGCTTCCGGAGAGCTGTGTAATCCGAGAGAGCTCCCGGGGAATATC 543
DB 556 GACCTCTGGACCTGCTTCCGGAGAGCTGTGTAATCCGAGAGAGCTCCCGGGGAATATC 615
QY 544 CTGATTCGGGTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGAGCGGCATTC 603
DB 616 CTGATTCGGGTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGAGCGGCATTC 675
QY 604 CTGCCAGCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCCCGC 663
DB 676 CTGCCAGCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCCCGC 735
QY 664 AGTCATCTGTGCTCTGCTCAAGTGGACCGAGCGGAAAGTGTGGGCCCAATTCGCGCAG 723
DB 736 AGTCATCTGTGCTCTGCTCAAGTGGACCGAGCGGAAAGTGTGGGCCCAATTCGCGCAG 795
QY 724 CGAGAGGGAACCTCTACCTGATTGACTTGGCTCGGTCAGAGGACAAACCGCGCACAGGC 783
DB 796 CGAGAGGGAACCTCTACCTGATTGACTTGGCTCGGTCAGAGGACAAACCGCGCACAGGC 855
QY 784 AACAGGGCTTCGGCTAAAGAGAGATGGAGGCCATCAACACCTCTGTTGTCTCGGC 843
DB 856 AACAGGGCTTCGGCTAAAGAGAGATGGAGGCCATCAACACCTCTGTTGTCTCGGC 915
QY 844 AAGTGTAGATGGCTGAATCAGGGCTCTGCTGTACTTATCGGAGCAGCAAGCTC 903
DB 916 AAGTGTAGATGGCTGAATCAGGGCTCTGCTGTACTTATCGGAGCAGCAAGCTC 975
QY 904 ACTCGCTATTGAGGACTCTCTGGGTGGCTCAGCCCAAGATATCTTATTCGCAACT 963
DB 976 ACTCGCTATTGAGGACTCTCTGGGTGGCTCAGCCCAAGATATCTTATTCGCAACT 1035
QY 964 GCCCTGAGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGGTC 1023
```

```
DB 1036 GCCCTGAGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGGTC 1095
QY 1024 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTGCAGCTCATGCTTGGGACCT 1083
DB 1096 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTGCAGCTCATGCTTGGGACCT 1155
QY 1084 GTTAAGCTGTCTCAGAAAGAAATTCCTTGGTCCACAGAGGCAAGAGAGCCCGAGGCGCT 1143
DB 1156 GTTAAGCTGTCTCAGAAAGAAATTCCTTGGTCCACAGAGGCAAGAGAGCCCGAGGCGCT 1215
QY 1144 GAGAAAGAGGATTTGGAGCCCTTGAGCCCATGAGCTCCAGCTCTGCTTCCAGAAA 1203
DB 1216 GAGAAAGAGGATTTGGAGCCCTTGAGCCCATGAGCTCCAGCTCTGCTTCCAGAAA 1275
QY 1204 CTCAGCCCCCTACAGAGACTTAAGCAGCATGAGACCCGCGCATGCTGGAGCGCTCTCTCAGC 1263
DB 1276 CTCAGCCCCCTACAGAGACTTAAGCAGCATGAGACCCGCGCATGCTGGAGCGCTCTCTCAGC 1335
QY 1264 TTGGACCTGCTGCTTGGCTCCAGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1323
DB 1336 TTGGACCTGCTGCTTGGCTCCAGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1395
QY 1324 CGAGAGCGGATGCTGCTTAATGAACAGTAGAAGAGAGACCTTAGAGATTGAGAGGCTT 1383
DB 1396 CGAGAGCGGATGCTGCTTAATGAACAGTAGAAGAGAGACCTTAGAGATTGAGAGGCTT 1455
QY 1384 AAGCAGAGCAAAAAGAACTGGAGGCCAAGATGTTGGCCCCAGAGGCTGAGAGAAAGAG 1443
DB 1456 AAGCAGAGCAAAAAGAACTGGAGGCCAAGATGTTGGCCCCAGAGGCTGAGAGAAAGAG 1515
QY 1444 AACATTGTCCCAATG 1461
DB 1516 AACATTGTCCCAATG 1533
```

RESULT 5

```
AR304057 LOCUS AR304057 2097 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 34 from patent US 6544766.
ACCESSION AR304057
VERSION AR304057.1 GI:31692955
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2097)
AUTHORS Beraud,C., Ohashi,C., Sakowicz,R., Vaiberg,E., Wood,K. and Yu,M.
TITLE Human kinesins and methods of producing and purifying human
kinesins
JOURNAL Patent: US 6544766-A 34 08-APR-2003;
FEATURES
source Location/Qualifiers
1..2097
/mol_type="genomic DNA"
```

ORIGIN

```
Query Match 99.6%; Score 1458; DB 6; Length 2097;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 GGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGGTAAGG 63
DB 100 GGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGGTAAGG 159
QY 64 GTGGCTGTGCACTCGGCCCAATTTGTGATGGAACAGCGGAGCAAGTATGATCCCCCTGT 123
DB 160 GTGGCTGTGCACTCGGCCCAATTTGTGATGGAACAGCGGAGCAAGTATGATCCCCCTGT 219
QY 124 GTGGGGGCATGGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACCAACAGAGACT 183
DB 220 GTGGGGGCATGGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACCAACAGAGACT 279
```

	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	1479	1480	1481	1482	1483	1484	1485	1486	1487	1488	1489	1490	1491	1492	1493	1494	1495	1496	1497	1498	1499	1500	1501	1502	1503	1504	1505	1506	1507	1508	1509	1510	1511	1512	1513	1514	1515	1516	1517	1518	1519	1520	1521	1522	1523	1524	1525	1526	1527	1528	1529	1530	1531	1532	1533	1534	1535	1536	1537	1538	1539	1540	1541	1542	1543	1544	1545	1546	1547	1548	1549	1550	1551	1552	1553	1554	1555	1556	1557	1558	1559	1560	1561	1562	1563	1564	1565	1566	1567	1568	1569	1570	1571	1572	1573	1574	1575	1576	1577	1578	1579	1580	1581	1582	1583	1584	1585	1586	1587	1588	1589	1590	1591	1592	1593	1594	1595	1596	1597	1598	1599	1600	1601	1602	1603	1604	1605	1606	1607	1608	1609	1610	1611	1612	1613	1614	1615	1616	1617	1618	1619	1620	1621	1622	1623	1624	1625	1626	1627	1628	1629	1630	1631	1632	1633	1634	1635	1636	1637	1638	1639	1640	1641	1642	1643	1644	1645	1646	1647	1648	1649	1650	1651	1652	1653	1654	1655	1656	1657	1658	1659	1660	1661	1662	1663	1664	1665	1666	1667	1668	1669	1670	1671	1672	1673	1674	1675	1676	1677	1678	1679	1680	1681	1682	1683	1684	1685	1686	1687	1688	1689	1690	1691	1692	1693	1694	1695	1696	1697	1698	1699	1700	1701	1702	1703	1704	1705	1706	1707	1708	1709	1710	1711	1712	1713	1714	1715	1716	1717	1718	1719	1720	1721	1722	1723	1724	1725	1726	1727	1728	1729	1730	1731	1732	1733	1734	1735	1736	1737	1738	1739	1740	1741	1742	1743	1744	1745	1746	1747	1748	1749	1750	1751	1752	1753	1754	1755	1756	1757	1758	1759	1760	1761	1762	1763	1764	1765	1766	1767
--	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

RESULT 6

Source

gene

CDS

```

/codon_start=1
/product="kinesin-like DNA binding protein"

```

```

/db_xref="GT:1549443"
/translation="RAAGSCQRRREMAAASAAAISGAGCRLSKIGATRRPPARV
RVARLURPVDTGACASDPVCRGNDSCSEIANRWNRHETLKVQDFAGBGRSTOOD
IYAGSVQTLRHLLEQONAVIAYGPTGAGKTHMLGSPQPGVILPALMDLQJLRE
EAGAEFPWALSTVMSYIEVKEJLDLPASGDVIRECPGONLIPGLMSQKPSISF
ADPERFLPASNRVGTALNRNRSRHVILVKVQDERLAPFPQREKGUJLDILA
GSEDRNTGKGLRKESGAINTSUFLGVKVDVADNGLCPVRPVRSKTLRLQDSJG
GSAHSILIANIAPERFYLDTPVSAANFAARKEVINRPTNESIOPHALCPGSLQSOKE
LILPPKAKRAGPGEESIEGPPMPAASASQKLSPIQLKISMDPAMLERLLSLDRLL
ASOGCAPLSTSPKRRRMVLMVEEKDLEIRKTKQKELAKMLAQAKAEKKNHC
PTMLRLSHRTVTGAKPLKXAVVMPQLIQQEAPASNAEHLKNGKGNKXLSBAL
EPKEACDWELQI5PEIILAGKQILDLNNEGARDLSQRIGPKXQQLI1VSWREL

```

HGPPSQVEDLSEVITGKQMESFLKANILGLAAGQRCAS*

ORIGIN	Query Match	Best Local Similarity	Score	1458;	DB 9;	Length	2097;	Mismatches	0;	Indels	0;	Gaps	0;
QY	4	GGTGGCTGCGCTAAGCAAGATTGGAGCTACTGTCGTCCACCTCCAGCTCGGTAAAG	63										
Db	100	GGTGGCTGCGCTAAGCAAGATTGGAGCTACTGTCGTCCACCTCCAGCTCGGTAAAG	159										
QY	64	GTGGCTGCGCACTGCGGCCAATTTGTGATGGAATGCTTAAGCAAGCAAGTATCCCTCTGT	123										
Db	160	GTGGCTGCGCACTGCGGCCAATTTGTGATGGAATGCTTAAGCAAGCAAGTATCCCTCTGT	219										
QY	124	GTGGGGGCGATGCGAGCTGCTCTAGAGATTGCTTAAGCAAGCAAGTATCCCTCTGT	183										
Db	220	GTGGGGGCGATGCGAGCTGCTCTAGAGATTGCTTAAGCAAGCAAGTATCCCTCTGT	279										
QY	184	CTCAAAATACCAATGATGCTTCTTAAGGAGGAGTACTCAGCAGGACATCTATGCA	243										
Db	280	CTCAAAATACCAATGATGCTTCTTAAGGAGGAGTACTCAGCAGGACATCTATGCA	339										
QY	244	GGTTGAGTGCAGCCCATCTAAGCACTTGTCTGGAAGGCGAGATGCCAGTGTGCTGCC	303										
Db	340	GGTTGAGTGCAGCCCATCTAAGCACTTGTCTGGAAGGCGAGATGCCAGTGTGCTGCC	399										
QY	304	TATGGAACCCACAGGAGCTGGGAAGCAGCACAATGCTGGGCGAGGAGGCTGGG	363										
Db	400	TATGGAACCCACAGGAGCTGGGAAGCAGCACAATGCTGGGCGAGGAGGCTGGG	459										
QY	364	GTGATCCCGGGGCTCTCATGGAGCTCTCGAGCTCACAAGGAGGAGGCTGGG	423										
Db	460	GTGATCCCGGGGCTCTCATGGAGCTCTCGAGCTCACAAGGAGGAGGCTGGG	519										
QY	424	CGGCCATGGGCGCTTCTGTACCATGCTTACCTAGAGATCTACAGGAGAGGATTA	483										
Db	520	CGGCCATGGGCGCTTCTGTACCATGCTTACCTAGAGATCTACAGGAGAGGATTA	579										
QY	484	GACCTCTGGACCTGCTTGGGAGAGCTGTGTAATCCGAGAGACTGCGGGGAATATC	543										
Db	580	GACCTCTGGACCTGCTTGGGAGAGCTGTGTAATCCGAGAGACTGCGGGGAATATC	639										
QY	544	CTGATTCGGGTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGAGCGGCATTC	603										
Db	640	CTGATTCGGGTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGAGCGGCATTC	699										
QY	604	CTGCCAGCAGTCAAAATCGAATGCTAGGAGCCACCCGGCTCAACAGCGCTCTCCCGC	663										
Db	700	CTGCCAGCAGTCAAAATCGAATGCTAGGAGCCACCCGGCTCAACAGCGCTCTCCCGC	759										
QY	664	AGTCATGCTGCTCTGCTCAAGGTGAGCCAGCGGGAAGCTTTGGCCCAATTTGCGCAG	723										
Db	760	AGTCATGCTGCTCTGCTCAAGGTGAGCCAGCGGGAAGCTTTGGCCCAATTTGCGCAG	819										
QY	724	CGAGAGGGAATCTACCTGATTTGACTTTGGCTGGGTCAAGGAGCAACCCGCGCAGAGC	783										
Db	820	CGAGAGGGAATCTACCTGATTTGACTTTGGCTGGGTCAAGGAGCAACCCGCGCAGAGC	879										
QY	784	AACAGGGCTTGGCTTAAAGAGTGGAGCCATCAACCTCCCTGTTTGTCTGGC	843										
Db	880	AACAGGGCTTGGCTTAAAGAGTGGAGCCATCAACCTCCCTGTTTGTCTGGC	939										
QY	844	AAAGTGTAGATGCGTGAATCAGGGCTCCCTGCTGCTTACTTATCGGACAGCAAGCTC	903										
Db	940	AAAGTGTAGATGCGTGAATCAGGGCTCCCTGCTGCTTACTTATCGGACAGCAAGCTC	999										
QY	904	ACTGCGCTATTGAGGAGCTCTCGGTGGGTCCAGCCCAAGTATCTTATTTGCCAATTT	963										
Db	1000	ACTGCGCTATTGAGGAGCTCTCGGTGGGTCCAGCCCAAGTATCTTATTTGCCAATTT	1059										
QY	964	GCCCTCGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGTCGAGGTC	1023										

RESULT 7	BC028155	2117 bp	mRNA	linear	PRI 29-JUN-2004
LOCUS	BC028155	2117 bp	mRNA	linear	PRI 29-JUN-2004
DEFINITION	Homo sapiens kinesin family member 22, mRNA (cdna clone MGC:40049 IMAGE:5241557), complete cds.				
ACCESSION	BC028155				
VERSION	BC028155.1	GI:20380446			
KEYWORDS	MGC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 2117)				
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buotow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Spatlet, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 2117)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				

JOURNAL	Submitted (08-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA			
REMARK COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgabbs@mail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc.mgc@nih.gov Akhtar, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Manduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.			
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILNL at: http://image.lnl.gov Series: IRAK Plate: 62 Row: 1 Column: 3 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein. Location/Qualifiers 1. .2117 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:40049 IMAGE:5241557" /tissue_type="Brain, fetal, whole pooled" /clone_lib="NIH_MGC_121" /lab_host="DH10B" /notes="Vector: pCMV-SPORT6" 1. .2117 /gene="KIF22" /notes="synonyms: KID, OBP, OBP-1, OBP-2, KNSL4" /db_xref="LocusID:3835" /db_xref="MIM:603213" 6. .2003 /gene="KIF22" /codon_start=1 /product="kinesin family member 22" /protein_id="AAH28155.1" /db_xref="GI:20380447" /db_xref="LocusID:3835" /db_xref="MIM:603213" translation="MAAGGTTQRRREMAAASAAISGACRLSKIGATRRPPPARV RVAVLPFFVDTAGADPPCVRGMDSCLEIANWRNHOETLVQDFADYGERSTQD ITYAGVDPILRHLEGQNSVLAAYPTGAGKTHMLGSPQPGVIPALMDLQLQTR EGAEGRWALSVNYSLEYOEKVLDDLPASGDLVIREDCRGNILIPGLSQPISSF ADPFHPLFASNRTVGATRLNQRSSHAVLLVKKVDQRLAPFRQREKTLIDILA GSDNRTGNKGLKESGAINTSFLVGLKVDALNQGLPRVPRSKLRLQDSIG GSAHILIANIAPERRFYLDVTSALNPAARKEVINRPTFNESLQHPALGPVKLSQK LLGPPEAKRARGPEEIGSGPEPMAAPASOKLSPLQKLSMDPAMLERLLSLDRLL ASQSQGNPLISTPKRPMVIMKTVEEKDLEIRLTKQKELEAKMLAOKAEKENHC PTMURPLSHRTVQAKLKVAVMPLQIQEQASPNAEIHLKNKRKEKLESLDAL EPBEKEDCWELQISPELLAHGRQKILDLNLESARDLRSQRIQPKKALQIVGNREL HGPFPSQVEDLERVEGITGKQMESFLKANIILGLAAGQRCCGAS"			
ORIGIN	Query Match 99.6%; Score 1458; DB 9; Length 2117; Best Local Similarity 100.0%; Pred. No. 0; Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	4	GGTCGCTGCTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCAGCTCCGCTAAGG	63	

Db 275 CTCAATACCAAGTTGATGCTCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 334
 Qy 244 GGTTCAGTGCAGCCCATCTAAGGACCTTGTCTGGAAGGCGAGAAATGCCAGTGTGCTTGGC 303
 Db 335 GGTTCAGTGCAGCCCATCTAAGGACCTTGTCTGGAAGGCGAGAAATGCCAGTGTGCTTGGC 394
 Qy 304 TATGACCCACAGGAGCTGGGAGAGCACAACAATGCTGGGAGCCACAGACCACTTGGG 363
 Db 395 TATGACCCACAGGAGCTGGGAGAGCACAACAATGCTGGGAGCCACAGACCACTTGGG 454
 Qy 364 GTGATCCCGGGCTCTCATGACCTCTGAGCTCACAAGGAGAGGGTGGCAGGGC 423
 Db 455 GTGATCCCGGGCTCTCATGACCTCTGAGCTCACAAGGAGAGGGTGGCAGGGC 514
 Qy 424 CGGCCATGGGCCCTTCTGTCTACCATGCTTCTACCTAGAGATCTACAGGAGAAAGTATTA 483
 Db 515 CGGCCATGGGCCCTTCTGTCTACCATGCTTCTACCTAGAGATCTACAGGAGAAAGTATTA 574
 Qy 484 GACCTCTGGACCTGCTTGGGAGACCTGTAATCCGAGAGACTGCGGGGGAATATC 543
 Db 575 GACCTCTGGACCTGCTTGGGAGACCTGTAATCCGAGAGACTGCGGGGGAATATC 634
 Qy 544 CTGATCCGGGCTCTCCAGAGCCCATCAGTGTCTGTCTGATTTTGGCGGCACCTTC 603
 Db 635 CTGATCCGGGCTCTCTCCAGAGCCCATCAGTGTCTGTCTGATTTTGGCGGCACCTTC 694
 Qy 604 CTGCCAGCCAGTCGAATCGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGC 663
 Db 695 CTGCCAGCCAGTCGAATCGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGC 754
 Qy 664 AGTCATGCTGTCTCTGCTCAAGTGGACCGAGGGAACGTTTGGCCCATTTGCCAG 723
 Db 755 AGTCATGCTGTCTCTGCTCAAGTGGACCGAGGGAACGTTTGGCCCATTTGCCAG 814
 Qy 724 CGAGAGGGAACCTCTACCTGATTTGCTGTGGTCTCAGAGCAACCGGCGCACAGGC 783
 Db 815 CGAGAGGGAACCTCTACCTGATTTGCTGTGGTCTCAGAGCAACCGGCGCACAGGC 874
 Qy 784 AACAGGGCTCTCGGCTAAAGAGTGGAGGCATCAACACCTCTGTTTCTCTGGGC 843
 Db 875 AACAGGGCTCTCGGCTAAAGAGTGGAGGCATCAACACCTCTGTTTCTCTGGGC 934
 Qy 844 AAGTGTAGATCGCTGATCAGGCTCTCTGCTGCTACCTTATCGGACAGCAGCTC 903
 Db 935 AAGTGTAGATCGCTGATCAGGCTCTCTGCTGCTACCTTATCGGACAGCAGCTC 994
 Qy 904 ACTCGCTATTGAGGACTCTCTGGTGGCTCAGGCCACAGATCTCTTATGCCAACATT 963
 Db 995 ACTCGCTATTGAGGACTCTCTGGTGGCTCAGGCCACAGATCTCTTATGCCAACATT 1054
 Qy 964 GCCCTGAGAGAGCTTTACCTAGACACAGTCTCGGACTCACTTTCTGCGAGGCTC 1023
 Db 1055 GCCCTGAGAGAGCTTTACCTAGACACAGTCTCGGACTCACTTTCTGCGAGGCTC 1114
 Qy 1024 AAGAGGTGATCAATCGGCTTTTACCAATGAGGCTTGAGGCTCATGCTTGGACCT 1083
 Db 1115 AAGAGGTGATCAATCGGCTTTTACCAATGAGGCTTGAGGCTCATGCTTGGACCT 1174
 Qy 1084 GTTAAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGCAAGAGAGCCGAGGCTC 1143
 Db 1175 GTTAAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGCAAGAGAGCCGAGGCTC 1234
 Qy 1144 GAGAAAGAGAGATTGGAGCCCTGAGCCCATGGCAGCTCCAGCTTCTGCTCCAGAAA 1203
 Db 1235 GAGAAAGAGAGATTGGAGCCCTGAGCCCATGGCAGCTCCAGCTTCTGCTCCAGAAA 1294
 Qy 1204 CTCAGCCCTTACAGAGAGCTTAAGCAGCATGGACCCGCGCCATGCTGGAGGCTCTCAGC 1263
 Db 1295 CTCAGCCCTTACAGAGAGCTTAAGCAGCATGGACCCGCGCCATGCTGGAGGCTCTCAGC 1354
 Qy 1264 TTGACCGCTGTGCTTCTCCAGGAGGAGCCAGGGGCGCCCTCTGTTGAGTACCCCAAG 1323
 Db 1355 TTGACCGCTGTGCTTCTCCAGGAGGAGCCAGGGGCGCCCTCTGTTGAGTACCCCAAG 1414

Qy 1324 CGAGAGCGGATGCTGTCTAATGAGACAGTAGAGAGAGGACCTAGAGATTGAGAGCTT 1383
 Db 1415 CGAGAGCGGATGCTGTCTAATGAGACAGTAGAGAGAGGACCTAGAGATTGAGAGCTT 1474
 Qy 1384 AAGACGAAGCAAAAGAACTGGAGGCAAGATGTTGGCCAGAGGCTGAGGAAAAGGAG 1443
 Db 1475 AAGACGAAGCAAAAGAACTGGAGGCAAGATGTTGGCCAGAGGCTGAGGAAAAGGAG 1534
 Qy 1444 AACCATTTGCCCAATG 1461
 Db 1535 AACCATTTGCCCAATG 1552

RESULT 9
 BC003427 2086 bp mRNA linear ROD 29-JUN-2004
 LOCUS Mus musculus kinesin family member 22, mRNA (cDNA clone MGC:6456
 DEFINITION IMAGE:2615715), complete cds.

ACCESSION BC003427
 VERSION BC003427.1 GI:13097359
 KEYWORDS MGC.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 2086)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Mallek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Vialla, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Pahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932
 2 (bases 1 to 2086)

Strausberg, R.
 Direct Submission

Submitted (20-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs@mail.nih.gov

Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loubegeed, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

TITLE
 JOURNAL
 PUBMED
 REFERENCE

AUTHORS
 TITLE
 JOURNAL

REMARK
 COMMENT

Series: IRAK Plate: 5 Row: m Column: 7
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21704181.
 Location/Qualifiers
 1. .2086
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="MGC:6456 IMAGE:2615715"
 /tissue type="mammary tumor, C3(1)-Tag model, infiltrating ductal carcinoma, 5 month old virgin mouse."
 /clone_lib="NCI CGAP_Mam6"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 1. .2086
 /gene="Kif22"
 /note="synonym: Kid"
 /db_xref="LocusID:110033"
 /db_xref="MGI:109233"
 20. .2002
 /gene="Kif22"
 /codon_start=1
 /product="kinesin family member 22"
 /protein_id="AAH03427.1"
 /db_xref="GI:13097360"
 /db_xref="LocusID:110033"
 /db_xref="MGI:109233"
 /translation="MSIRAKTCPORREMASATSPGRCVSGKIGRPPILRVAVR
 LRPMDETAKELPCVRAIDSCSLEVANKKYOETLUYQDAPIGSKSQEYVGS
 VQILRLHLLGQNSVLAIGTGTAGKTHLMGSPQGVIPRALMDLLQARLESABG
 RPDVMSYLEYIQKVLDDLPASGLVIREDCRNILIPGLTKPITSFSDPQ
 HFLPSNRVAVPRLNQRSRSHVLLVKVDORERTPRQREGLYLDLAGSEDN
 RRTGNQIRLKESGAINSLFVLGVVDALNOGLPRIYPRDSKLTRLOSLGSSAHS
 ILIANIAPERFYQDTISALNFTARSKEVINRPTNESLOPHALA PVKLSQKELLGRS
 EAKKAPPEESGPSSTAPASQKLSLLQKLSNMDPAMLENLUSMERLLGSSGS
 QGPTLLTPRRERVMKTVREKNLEIRLKKOKLEAKVLAQEPDPREKENTPTI
 LOPPASYSYGVAPLKKAVMPQRIQKRESSNOIQLLKKPKRKLPEPSEAVEK
 DEYWEVQISPELLAHGRKLLDLNREGSRELRSRLQIGQKKAQLIVGRELHGPS
 EVEDLEQVEGISGQVESFLKANLLSLAASQHSGPS"

FEATURES

source

gene

CDS

ORIGIN

Query Match 72.9%; Score 1067.8; DB 10; Length 2086;
 Best Local Similarity 84.1%; Pred. No. 2.2e-274;
 Matches 1204; Conservative 0; Mismatches 227; Indels 0; Gaps 0;
 QY 6 TCGCTGTCGCTAGCAGATTTGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGGCT 65
 DB 82 TGGCCGCTGTGTAAGCAAGAGAGGGCTTGGCCGCGTCCGCTCTAGCTCGAGTAAGGGT 141
 QY 66 GGCTGTGCGACTGGCGGCATTGTGGATGGAACAGCGGAGCAAGTATCCCCCTGTGT 125
 DB 142 GGCTGTGCGGCTGGCCGGTTTATGGATGGAGACAGAGCGGAGAGCTCCCTGTGT 201
 QY 126 CGCGGGCATGAGACAGCTGCTCTTAGAGATTGCTTAAGTGGAGAACCCAGGAGACTCT 185
 DB 202 CCGAGCCATAGACAGCTGCTCTTGAAGTGGCTAACTGGAAGAAATACCCAGGAGACGCT 261
 QY 186 CAATACCATGTTGATCCCTCTATGGGAGAGAGAGTACTCAGCAGACATCTATGACGG 245
 DB 262 CAATATACAGTTTGTATCCCTCTATGCGGAGAGAGCACTCAGCAGGAAGTCTATGTAGG 321
 QY 246 TTCAAGTCAGCCATCTTAAGGCACTTGTGGAGGGCAGAAATGCCAGTGTGCTTGCCTA 305
 DB 322 TTCAAGTCAGCCATCTTAAGGCACTTGTGGAGGGCAGAAATGCCAGTGTGCTTGCCTA 381
 QY 306 TGGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGCGCCAGCAGCAACTGGGGT 365
 DB 382 TGGGCTTACTTGGGCGAGGAGACACACATGCTGGGAGCGCCAGCAGCAACTGGAGT 441
 QY 366 GATCCCGCGGGCTCTCATGAGCTCTCTGAGCTCACAAGGAGAGAGGCTCCCGAGGCGG 425
 DB 442 GATTCTCGGGCTCTCATGAGCTCTCTGAGCTCACAAGGAGAGAGTCTCGAGGAGCG 501

QY 426 GCCATGGGCCCTTTCTGTCAACATGTTTACTTAGAGATCTACACGAGAGAGGTATTAGA 485
 DB 502 GCCATGGAGCGTTTCTGTGCTAATGCTCTATTAGAGATCTACACGAGAAAGGTATTAGA 561
 QY 486 CCTCTGGACCCCTGCTTCGGGAGACCTGGTAATCCGAGAAAGTCTGCCGGGGGAATATCT 545
 DB 562 CCTCTGGATCTGCTCATCAGGAGACCTCGTATCCGGAAGACTGCCGAGGGAACATCT 621
 QY 546 GATTCCGGGCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTCTCT 605
 DB 622 GATCCAGGCTCTACACAGAGCCCATCAGAGTCTTCTGACTTCGAGAGCAGCTTCTCT 681
 QY 606 GCCAGCCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACAGAGCTCTCTCCGCGAG 665
 DB 682 TCCAGCCAGTAGAATCGAGCTGTAGGAGCCACCGGCTTAACAGCGCTCTTCCCGTAG 741
 QY 666 TCATGCTGTGCTCTGCTCAAGGTGACACGAGCGGGAAGCTTTGGCCCTCATTTGCGCAGC 725
 DB 742 TCACGAGTGTCTTGGTCAAGGTAGATCAGCGTGAACGTTTGTACTTCCATTTCCGCGC 801
 QY 726 AGAGGGAAGAACTCTACCTGATTGACTTTGGCTGGGTGAGGACAAACCGGCGCACAGGCAA 785
 DB 802 GGAAGGAAGAACTCTACCTTATTGATTGGCTGGTTCAGAGGACAAACCGTTCACAGGCAA 861
 QY 786 CAAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTTGTCTTGGGCAA 845
 DB 862 CCAGGGCATTTGGCTCAAGAGGAGCGGAGCCATCAACACCTCTCTTGTACTTGGGCAA 921
 QY 846 AGTGATAGTGGCTGTAATCAGGCGCTCTCTGCTGTTTACCTTATCGGACAGCAAGCTCAC 905
 DB 922 AGTGTGGATGCTAAACCCAGGCGCTCTCTCGCATACCATACCGGAGACAGCAAGCTCAC 981
 QY 906 TCGCTATTTCGAGGACTCTCTGSGTGGCTCAGCCCAAGTATCTTATTGCAACATTCG 965
 DB 982 TCGTCTGCTCGAGGACTCTCTGSGAGGCTCAGCTCATAGCATCTCTTATGCAACATTCG 1041
 QY 966 CCCTGAGAGAGCTTCTTACCTAGACAGTCTCGGCACTCAACTTGTGTCGACGCTCAA 1025
 DB 1042 TCTGGAAGAGCTTTTACCAGGATCAATTTAGCATTTGACTTCTACTGCTAGTCCAA 1101
 QY 1026 GGAGGTGATCAATCGGCTTTTACCAGGAGCTGAGGCTCTGAGCTCTATGCTTGGAGCTGT 1085
 DB 1102 GGAGGTGATTAACCGGCTTTTACCAGGAGCTGAGGCTCTGAGCTCTATGCTTGGAGCTGT 1161
 QY 1086 TAAGCTGTCTCAGAAAGAAATTTGCTTGTCCACAGAGGCAAGAGAGCCCGAGGCGCTTGA 1145
 DB 1162 TAAGCTGTCTCAGAAAGAAATTTGCTTGTCCCGCTCAGAGGCAAGAGAGCCCGCTTGA 1221
 QY 1146 GGAAGAGGAGATTTGGGAGCCCTGAGCCCATGGGAGCTTCCAGCTCTGCTTCCCGAGAACT 1205
 DB 1222 AGAAGATCAATCTGGGAGCCCTGAGTCTAGAGAGCTTCCAGCTCTGCTTCCCGAGAACT 1281
 QY 1206 CAGCCCCCTTACAGAAAGCTAAGCAGCATGGACCCCGGCAATCTGAGAGCGCTCTCTCAGCTT 1265
 DB 1282 AAGCTCTTACAGAAAGCTAAGCAATATGGATCCAGCCATGTTGGAGAACTCTCTGAGTAT 1341
 QY 1266 GGACCGTCTGCTTCCCTCCAGGAGCCAGGGGGCCCTCTGTTGAGTATACCCCAAGAGCG 1325
 DB 1342 GGAACGTTTCTGCGGTTTCCAGGGAAGCCAAAGGAGACCCCTTACTGTAATACCCCAAGAGCG 1401
 QY 1326 AGAGCGGATGCTGCTAATGAAGACAGTAGAAGAGAGAGGACCTAGAGATTTCAGAGGCTTAA 1385
 DB 1402 AGAACGATGCTGCTCATGAAGACAGTGGAGGAAAGAACTTAGAGATTTCAGAGGCTTAA 1461
 QY 1386 GACCAAGCAAAAAGAACTGGAGGCGCAAGATGTTGGCCCGAGAGGCTCAGGA 1436
 DB 1462 GATGAAGCAAAAAGAACTGGAAGCCAAAGTGTGTGGCCCGAGGAGCCCGGA 1512

RESULT 10
 AR210054
 LOCUS

AR210054 1041 bp DNA linear PAT 20-JUN-2002

```

DEFINITION Sequence 5 from patent US 6387644.
ACCESSION AR210054
VERSION AR210054.1 GI:21512185
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1041)
AUTHORS Beraud, C.
TITLE Motor proteins and methods for their use
JOURNAL Patent: US 6387644-A 5 14-MAY-2002;
FEATURES
    Location/Qualifiers
        1..1041
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN
    Query Match 70.9%; Score 1038.4; DB 6; Length 1041;
    Best Local Similarity 99.9%; Pred. No. 1.6e-266;
    Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTGCTCACTCCAGCTCGGTA 60
Db 1 ATGGGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTGCTCACTCCAGCTCGGTA 60
QY 61 AGGGTGGCTGTGGCACTGGCGCCATTGTGGTGAACACAGCGGGAGCAAGTATCCCCC 120
Db 61 AGGGTGGCTGTGGCACTGGCGCCATTGTGGTGAACACAGCGGGAGCAAGTATCCCCC 120
QY 121 TGTGTGGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACACCAAGAG 180
Db 121 TGTGTGGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACACCAAGAG 180
QY 181 ACTCTCAATACCAATTTGATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTAT 240
Db 181 ACTCTCAATACCAATTTGATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTAT 240
QY 241 GCAGGTTCACTGACAGCCCATCTTAAGGCACTTGTGGAAGGAGGAGATGCCAGTGT 300
Db 241 GCAGGTTCACTGACAGCCCATCTTAAGGCACTTGTGGAAGGAGGAGATGCCAGTGT 300
QY 301 GCCTATGGACCCACAGAGCTGGGAAGAGACACACAACTGTGGGAGCCACAGAGCACT 360
Db 301 GCCTATGGACCCACAGAGCTGGGAAGAGACACACAACTGTGGGAGCCACAGAGCACT 360
QY 361 GGGGTGATCCCGGGGCTCTCATGGAACCTCTGTCAGCTCACAAGGAGGAGGTGCCGAG 420
Db 361 GGGGTGATCCCGGGGCTCTCATGGAACCTCTGTCAGCTCACAAGGAGGAGGTGCCGAG 420
QY 421 GGGGGGATCCCGGGGCTCTCATGGAACCTCTGTCAGCTCACAAGGAGGAGGTGCCGAG 480
Db 421 GGGGGGATCCCGGGGCTCTCATGGAACCTCTGTCAGCTCACAAGGAGGAGGTGCCGAG 480
QY 481 TTAGACTCTCTGACCTCTGTCGGGAGACCTGCTAATCCGAGAGACTGCCGGGGAAT 540
Db 481 TTAGACTCTCTGACCTCTGTCGGGAGACCTGCTAATCCGAGAGACTGCCGGGGAAT 540
QY 541 ATCCTGATCCGGGTCTCTCCAGAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCAC 600
Db 541 ATCCTGATCCGGGTCTCTCCAGAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCAC 600
QY 601 TTCTGCGCAGCTCGAATTCGAGTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCC 660
Db 601 TTCTGCGCAGCTCGAATTCGAGTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCC 660
QY 661 CGCAGTCTGCTGTGCTCTGGTCAAGGTGGACAGCGGGAAGCTTTGGCCCCCATTTCC 720
Db 661 CGCAGTCTGCTGTGCTCTGGTCAAGGTGGACAGCGGGAAGCTTTGGCCCCCATTTCC 720
QY 721 CAGCGAGAGGGAACCTCTA CTTGATTTGCTGGGTGAGTGTGGGTGAGTGTGGGTGAG 780
Db 721 CAGCGAGAGGGAACCTCTA CTTGATTTGCTGGGTGAGTGTGGGTGAGTGTGGGTGAG 780

```

```

QY 781 GGCAACAAGGGCTTTCGGCTAAAGAGAGAGTGGAGCCATCAACACCTCTCTGTTCTCTG 840
Db 781 GGCAACAAGGGCTTTCGGCTAAAGAGAGAGTGGAGCCATCAACACCTCTCTGTTCTCTG 840
QY 841 GGCAAAAGTGGTAGATGCGCTGAATCAGGGCCTCTCTGTTACCTTATCGGACAGCAAG 900
Db 841 GGCAAAAGTGGTAGATGCGCTGAATCAGGGCCTCTCTGTTACCTTATCGGACAGCAAG 900
QY 901 CTCACCTCGCTATTTCAGGACTCTCTGGTGGCTCAGCCACAGTATCTTATTCGCAAC 960
Db 901 CTCACCTCGCTATTTCAGGACTCTCTGGTGGCTCAGCCACAGTATCTTATTCGCAAC 960
QY 961 ATTGCCCCCTGAGAGACGCTTCTACCTAGACACAGTCTCGCACTCAACTTTGCTGCCAG 1020
Db 961 ATTGCCCCCTGAGAGACGCTTCTACCTAGACACAGTCTCGCACTCAACTTTGCTGCCAG 1020
QY 1021 TCCAAGGAGGTGATCAATTCG 1040
Db 1021 TCCAAGGAGGTGATCAATTCG 1040

RESULT 11
LOCUS AR210052 1115 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 1 from patent US 6387644.
ACCESSION AR210052
VERSION AR210052.1 GI:21512182
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1115)
AUTHORS Beraud, C.
TITLE Motor proteins and methods for their use
JOURNAL Patent: US 6387644-A 1 14-MAY-2002;
FEATURES
    Location/Qualifiers
        1..1115
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN
    Query Match 70.7%; Score 1035.4; DB 6; Length 1115;
    Best Local Similarity 99.9%; Pred. No. 1e-265;
    Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGTGCTGTGCTGCTAAGCAAGATTGGAGCTACTCGTCTGCTCACTCCAGCTCGGTAAGG 63
Db 78 GGTGCTGTGCTGCTAAGCAAGATTGGAGCTACTCGTCTGCTCACTCCAGCTCGGTAAGG 137
QY 64 GTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 123
Db 138 GTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 197
QY 124 GTGGGGGATGACGACAGCTGCTCTCTAGAGATTGCTAATCGAGGAACACCAAGAGACT 183
Db 198 GTGGGGGATGACGACAGCTGCTCTCTAGAGATTGCTAATCGAGGAACACCAAGAGACT 257
QY 184 CTCAAATACCAAGTTGATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTATGCA 243
Db 258 CTCAAATACCAAGTTGATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTATGCA 317
QY 244 GGTTCAGTGCAGCCCATCTTAAGGCACTTGTGGAAGGGCAGATGCCAGTGTGCTTGC 303
Db 318 GGTTCAGTGCAGCCCATCTTAAGGCACTTGTGGAAGGGCAGATGCCAGTGTGCTTGC 377
QY 304 TATGACCCACAGAGAGCTGGGAAGAGCAGCACAAATGCTGGGAGCCACAGAGCAACCTGG 363
Db 378 TATGACCCACAGAGAGCTGGGAAGAGCAGCACAAATGCTGGGAGCCACAGAGCAACCTGG 437
QY 364 GTGATCCCGGGGCTCTCATGAGCTCTCTGAGTCTCTGAGTCTCAAGGGAGGAGGTGCCAGGGC 423
Db 438 GTGATCCCGGGGCTCTCATGAGCTCTCTGAGTCTCTGAGTCTCAAGGGAGGAGGTGCCAGGGC 497

```

```
QY 424 CGGCCATGGGCCCTTCTGTGTCACGATGCTTACCTAGAGATCTACCGAGGAAGATATTA 483
Db 498 CGGCCATGGGCCCTTCTGTGTCACGATGCTTACCTAGAGATCTACCGAGGAAGATATTA 557
QY 484 GACCTCTCGGACCTGCTTCGCGAGACCTGCTATCCGAGAGACTGCGGGGGAATATC 543
Db 558 GACCTCTCGGACCTGCTTCGCGAGACCTGCTATCCGAGAGACTGCGGGGGAATATC 617
QY 544 CTGATTCGGGTCTCTCCAGAAAGCCCATCAGTAGCTTGTGCTGATTTTGAGCGGCACATTC 603
Db 618 CTGATTCGGGTCTCTCCAGAAAGCCCATCAGTAGCTTGTGCTGATTTTGAGCGGCACATTC 677
QY 604 CTGCAGCAGTCGAAATCGAGCTGTAGAGACACCGCGCTCAACAGCGCTCTCTCCGCG 663
Db 678 CTGCAGCAGTCGAAATCGAGCTGTAGAGACACCGCGCTCAACAGCGCTCTCTCCGCG 737
QY 664 AGTCATGCTGTGCTCTGCTCAAGTGGACGAGCGGGAACCTTTGGCCCAATTTGCGCAG 723
Db 738 AGTCATGCTGTGCTCTGCTCAAGTGGACGAGCGGGAACCTTTGGCCCAATTTGCGCAG 797
QY 724 CGAGAGGGAACCTCTACCTGATTGACTTGGCTGGGTGAGAGACAAACCGCGCGCACAGGC 783
Db 798 CGAGAGGGAACCTCTACCTGATTGACTTGGCTGGGTGAGAGACAAACCGCGCGCACAGGC 857
QY 784 AACAGGGCCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCCCTGTTTGTCTGGGC 843
Db 858 AACAGGGCCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCCCTGTTTGTCTGGGC 917
QY 844 AAAGTGTAGATGCGCTGAATCAGGCGCTCCCTGCTGCTACCTATCGGACAGCAAGCTC 903
Db 918 AAAGTGTAGATGCGCTGAATCAGGCGCTCCCTGCTGCTACCTATCGGACAGCAAGCTC 977
QY 904 ACTCGCTATTGACGAGCTCTCTCGGTGGCTCAGCCACAGATACCTTATTGCCAACATT 963
Db 978 ACTCGCTATTGACGAGCTCTCTCGGTGGCTCAGCCACAGATACCTTATTGCCAACATT 1037
QY 964 GCCCTGAGAGACGCTTCTACCTAGACACAGTCTCCGACATCAACTTGTGTCGAGTCC 1023
Db 1038 GCCCTGAGAGACGCTTCTACCTAGACACAGTCTCCGACATCAACTTGTGTCGAGTCC 1097
QY 1024 AAGGAGTGATCAATCG 1040
Db 1098 AAGGAGTGATCAATTG 1114

RESULT 12
AC101919/c
LOCUS AC101919 151041 bp DNA linear HTG 14-MAR-2003
DEFINITION Mus musculus clone RP24-158A6, WORKING DRAFT SEQUENCE, 4 unordered
pieces.
ACCESSION AC101919
VERSION AC101919.4 GI:28951338
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 151041)
Birren,B., Nusbbaum,C. and Lander,E.
Mus musculus, clone RP24-158A6
Unpublished
2 (bases 1 to 151041)
Birren,B., Linton,L., Nusbbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepeil,Y., Colangelo,M., Collins,S., Collimore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
```

```
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 151041)
Birren,B., Nusbbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepeil,Y.,
Collimore,A., Cooke,P., Cook,A., Corum,B., DeArellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hagaz,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Teafaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (14-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 14, 2003 this sequence version replaced gi:28631351.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L17745
Center clone name: 158 A.6
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 150558 bases at least Q40
Consensus quality: 150654 bases at least Q30
Consensus quality: 150687 bases at least Q20
Insert size: 151000; agarose-fp
Insert size: 150741; sum-of-contents
Quality coverage: 11.8 in Q20 bases; agarose-fp
Quality coverage: 11.8 in Q20 bases; sum-of-contents
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
* 33517: contig of 33517 bp in length
* 33518 33617: gap of 100 bp
* 33618 42736: contig of 9119 bp in length
```

[illegible]

Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Coyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karat, A., Karatas, A., Kells, C., LaRocque, K., Lanazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, J., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roberti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 164759)
 Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (17-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 164759)
 Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Submitted (23-APR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 23, 2004 this sequence version replaced gi:45504330.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence submissions@broad.mit.edu
 ----- Project Information
 Center project name: L17344
 Center clone name: 344_C18

FEATURES	source	Location/Qualifiers
repeat_region	1..164759	/organism="Mus musculus"
misc_feature	1..1655	/mol_type="genomic DNA"
		/db_xref="taxon:10090"
		/chromosome="1"
		/map="1"
		/clones="RP24-344C18"
		/clone_lib="RPC1-24 Male Mouse BAC"
		/rpt_family="L1_MM"
		/note="clone boundary"
		clone_end:SP6
		site:MboI
repeat_region	1767..1787	/rpt_family="AT_rich"
repeat_region	2115..2157	/rpt_family="AT_rich"
repeat_region	2236..2258	/rpt_family="AT_rich"
repeat_region	3252..3420	/rpt_family="AT_rich"
repeat_region	3464..3588	/rpt_family="Lx7"
repeat_region	3605..3666	/rpt_family="Lx7"
repeat_region	3673..3731	/rpt_family="CT-rich"
repeat_region	3732..3911	/rpt_family="(TTC)n"
repeat_region	4000..4180	/rpt_family="(TTC)n"
repeat_region	4202..4363	/rpt_family="CT-rich"
repeat_region	4529..4765	/rpt_family="Lx7"
repeat_region	7234..7442	/rpt_family="MTB"
repeat_region	8445..8509	/rpt_family="B3"
repeat_region		/rpt_family="T-rich"
repeat_region		complement (8512..8591)
repeat_region		9336..9632
repeat_region		/rpt_family="MLT1J"
repeat_region		complement (9800..11396)
repeat_region		/rpt_family="L1_MM"
repeat_region		complement (11397..17527)
repeat_region		/rpt_family="L1F"
repeat_region		complement (18235..18777)
repeat_region		/rpt_family="Lx"
repeat_region		18775..18890
repeat_region		/rpt_family="Lx"
repeat_region		19260..19949
repeat_region		/rpt_family="RMER19A"
repeat_region		20118..20297
repeat_region		/rpt_family="MLT1D"
repeat_region		complement (20386..20639)
repeat_region		/rpt_family="Lx2"
repeat_region		20634..20839
repeat_region		/rpt_family="Lx4"
repeat_region		20876..21793
repeat_region		/rpt_family="RLTR13D"

repeat_region	21799..21907	/rpt_family="RMR13A"	1492	GGACCTGCTTCGGGAGACCTGGTAATCCGAGAGAGACTCGCGGGGGAATATCTCGATTCC	551
repeat_region	21918..22433	/rpt_family="Lx4"	145633	GGATCCTGTCATCAGAAGACCTAGTGTTCGGAAGAGCTGACGAGGAAATCTCTGATTCC	145574
repeat_region	22591..22651	/rpt_family="(TG)n"	552	GGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTGATTTTGAAGCGGCACTTCTTCGCGAGC	611
repeat_region	22733..24486	/rpt_family="L1_MM"	145573	AGGCTCACAGAAAGCCCATCAGTAGTTTCTCTGACTTCGAGAAGCACTTCTCTCTAGT	145514
repeat_region	24494..24850	/rpt_family="L1_MM"	612	CAGTCGAAATCGGACTGTAGAGGCCACCGGCTCAACAGCGCTCTCCCGCAGTCAATGC	671
repeat_region	27139..27166	/rpt_family="L1_MM"	145513	CAGTAGAAATCGAGTTGTAGGAGCTACTCTGCTTTAACGAGCGCTCTTGCGCTAGTCAATGC	145454
repeat_region	27737..27834	/rpt_family="AT_rich"	672	TGTGCTCCTGCTCAAGTGGACAGCGGGAACGTTTGGCCCCATTTGCCCCATTTGCCCCAGAGAGG	731
repeat_region	28328..28506	/rpt_family="MLT1H"	145453	AGCGCTCTTGGTCAAGGTAGATCAGCATGAACGTTTGAATCCATTTCCGACAGTGGGAAGG	145394
repeat_region	28513..28541	/rpt_family="MLT1K"	732	AAAACCTTACCTGATTTGACTTTGGCTGGGTAGAGGCAACACC-----GGGCGACAGGCAACA	787
repeat_region	28813..28858	/rpt_family="AT_rich"	145393	AAAACCTTACCTTATTGATTTGGCTGGTTTCAGAGGCAACCGTCACTGTCGACAGGCAACC	145334
repeat_region	29627..30186	/rpt_family="AT_rich"	788	AGGCGCTTCGGCTAAAGAGAGTGGGCCATCAACACCTCCCTGTTTCTCTGGGCAAG	847
repeat_region	30465..30578	/rpt_family="L1_MM"	145333	AGGGCAITTTCTGCTCAAGAGAGTGGAGCCATCAACACCTCCCTCTTTGTACTGGGCAAG	145274
repeat_region	36749..36920	/rpt_family="T-rich"	848	TGCTAGATGCGCTGAATCAGGGCCCTCCCTCGTGTACTTATCGGACAGCAAGCTCACTC	907
repeat_region	37617..37751	/rpt_family="Lx2"	145273	TAGTGATGTCATTAACCAAGGCCCTCTCATATACCATACAGGACAGGAGGCTCACTC	145214
repeat_region	37958..37986	/rpt_family="(TA)n"	908	GCCTATTCCAGGACTCTCTGGGTGGCTCAGGCCACAGTATCTTATTTGCCAATTTGCC	967
repeat_region		/rpt_family="Lx4"	145213	GTCTGTGCAAGACTCTCTGAGAGGCTCAGCTCATAGCATCTCTATTGTCAGCAATTTGCTC	145154
repeat_region			968	CTGAGAGAGCTTCTACCTAGACAGTCTCGGCATCAACTTTGCTGCGCAGGCTCAAAGG	1027
repeat_region			145153	CTGAAGAGTGTTTTACCTGGATACATCTCAGCATTAACCTTCACTGCTAGTCCCAAGG	145094
repeat_region			1028	AGGTGATCAATCGGCTTTTACCAATGAGAGCTCAGCCCTCATGCCCTTGGGACCTGTTA	1087
repeat_region			145093	AGGTGATTAACCAAGGCTTTTACCAAGCGAGTGTACAACCTCAGGCTTGACACCTGTTA	145034
repeat_region			1088	AGCTGTCTCAGAAAGTGTCTTGTCACAGAGGCAAGAGAGCCGAGGCGCTGAGG	1147
repeat_region			145033	AGCTGTCTCAGAAAGTGTCTTGAGTTTATCAGAGGCAAGAAAGCCAAAGGCCCTCAAG	144974
repeat_region			1148	AAAGAGAGATTCGGAGCCCTGAGCCATGCGAGCTCCAGAGCTCTGCTCTCCAGAAACTCA	1207
repeat_region			144973	AAAGATCAACTGGAGCCCTGAGTCTTAGCATCTCCAGCATCTGCTTCCAGAAA-----	144918
repeat_region			1208	GGCCCCCTACAGAAAGCTAAGCAGCATGGACCCCGCCATGTGGAGCGGCTCTCTCAGTTGG	1267
repeat_region			144917	-----CTAAGCAATATGGATCTAGCCATGTGGGAAAGCTCTCTGAGTATGG	144872
repeat_region			1268	ACGCTCTGCTTCCAGGGAGCCAGGGGCGCCCTCTGTTGAGTACCCCAAAAGCGAG	1327
repeat_region			144871	AATGTTTGTGGTTTCAAGGGAGCCAGAGAACCTCTACTGAATACCCCAAA--CGAG	144813
repeat_region			1328	AGCGGATGTGCTTAATGAAGACAGTAGAAGAGAGGACCTTAGAGATTTGAGGCTTAAAG	1387
repeat_region			144812	AGCGAATGTGCTCATTAAGCGTGGAGGACAGATATTTGGAGATTTGAGAGGCTTAAAG	144753
repeat_region			1388	CGAAGCAAAAGAACTGGAGGCCAAGATGTTGGCCCGCAAGGC	1430
repeat_region			144752	TGAAGCAAAAGAACTGGAGGCCAGGTCGTGGTCCAGGAAGC	144710

RESULT 14
AC101908/c
LOCUS
DEFINITION
Mus musculus chromosome 1 clone RP24-275J1 map 1, *** SEQUENCING IN
PROGRESS ***, 8 unordered pieces.
AC101908
AC101908.4 GI:44199137
HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.

Query Match	60.2%; Score 881.8; DB 10; Length 164759;
Best Local Similarity	78.4%; Pred. No. 1.8e-224;
Matches 1116; Conservative	0; Mismatches 282; Indels 25; Gaps 4;
QY	12 TCGGCTAAGCAAGATTGGAGCTACTCGTCTGCACCTCCAGCTCGGCTAAGGTTGGCTGT 71
Db	146111 TTGTATAGCAAGTAGGATTTGGCTGTCTCTTCTAGCTTGAAGTGAAGGTGGCTAT 146052
QY	72 GCGACTCGCGCCATTTGTGGATGGAACAGCGGAGCAAGTGAATCCGCCCTGTGTGGGGG 131
Db	146051 ATAAATTGTGCTGTTTGTGGATGAAGACAGAAAGCAAGGAGCACCCCTGTGTACGAGC 145992
QY	132 CATGGACAGCTGCTCTTAGAGATTGCTTAAGTGGAGAAACACAGGAGACTCTCAAAATA 191
Db	145991 CATAGACAGCTGCTCTCTTTGAAGTGGCTTAACTGGAGGAAATACAGGAGACTATCAAAATA 145932
QY	192 CCAGTTTGTATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATCAGGTTTCAGT 251
Db	145931 TCAGTTTGTATGCTTCTTCAATGGAGAGAAAGACATCTAGTGGGACATCTATGTAGTTTAGT 145872
QY	252 GCAGCCCATCTTAAGGCACTTCTCTGGAAAGGCGAGAATGCCAGTGTGCTTCCCTATGACC 311
Db	145871 ACAGCCCATCTAAGGAACTTCTCTGGAAGGCGAGAATGCCAGTGTACTTGCCTATGACC 145812
QY	312 CACAGGAGCTGGGAAGACGACACAAATGCTGGGAGCCGACAGAGCAACTGGGTTGATCCC 371
Db	145811 GACTGGGGCAGGGAAGA--CACACAATGTTGGGAGCCGACAGCAGCTGGAGTGATTC 145754
QY	372 GCGGCTCTCATGGACCTCTCGAGCTCACAAGGAGGAGGGTCCGAGGCGCGGCCATG 431
Db	145753 TCAGACTCTCCAGACCTCTCTGCACTTAACAGGGAAGAGAACTCTGAGGCTGGCCATG 145694
QY	432 GGCCCTTTCTGTACCATGCTCTTACCTAGAGATCTTACCAGGAGAGGTATTAGACCTCTCT 491
Db	145693 GGACATTTCTGTGCTTGTGCTATTTAGATCTCTAAAGGAAATAGTATTAGACCTCTT 145634

SOURCE
ORGANISM

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 297639)

Birren, B., Nusbaum, C. and Lander, E.

Mus musculus chromosome 1, clone RP24-275J1

TITLE

Unpublished

REFERENCE
AUTHORS

2 (bases 1 to 297639)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B.,
Brown, A., Camarata, J., Campiano, A., Chang, J., Chazaro, B.,
Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooper, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kanat, A., Karatas, A., Kells, C., LaRoque, K.,
Lamaze, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McSwan, P., McKernan, K., McPheeters, R., Meldrim, J.,
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigliio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

TITLE

JOURNAL

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

3 (bases 1 to 297639)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V.,
Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J.,
Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B.,
Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L.,
Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D.,
Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafiz, N.,
Hegopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., Landers, T.,
Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R.,
Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
Zimmer, A. and Zody, M.

Direct Submission

TITLE

JOURNAL

Submitted (18-JUN-2004) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Feb 27, 2004 this sequence version replaced gi:31880232.

All repeats were identified using RepeatMasker:

Smith, A. F. A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: L17732

Center Clone name: 275_J_1

* NOTE: This is a 'working draft' sequence. It currently

* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 63747: contig of 63747 bp in length
63848 63847: gap of 100 bp
63848 73892: contig of 10045 bp in length
73893 73992: gap of 100 bp
73993 154450: contig of 80458 bp in length
154451 154550: gap of 100 bp
154551 167808: contig of 13258 bp in length
167809 167908: gap of 100 bp
167909 167911: contig of 8803 bp in length
167912 176811: gap of 100 bp
176812 233500: contig of 56689 bp in length
233501 233600: gap of 100 bp
233601 241222: contig of 7622 bp in length
241223 241322: gap of 100 bp
241323 297639: contig of 56317 bp in length.

FEATURES
source

1. 297639
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/map="1"
/clones="RP24-275J1"
/clone_lib="RPC1-24 Male Mouse BAC"

ORIGIN

Query Match	50.6%;	Score	740.6;	DB 2;	Length	297639;
Best Local Similarity	78.2%;	Pred. No.	1.1e-186;			
Matches	916;	Conservative	0;	Mismatches	249;	Gaps 2;

Qy	12	TCGGCTAAGCAAGATTGGAGCTACTCTGTCGTCACCTCCAGCTCGCGTAAGGGTGGCGT	71
Db	1179	TTGTATTAAGCAAGTAGGATTGGCTGCTGCTCTTAGCTGAGTAAGGGTGGCTAT	1120
Qy	72	CGCACTCGGCGCAATTTGTGGATGAAACAGCGGGAGCAAGTGTATCCCTCTGTGCGGGG	131
Db	1119	ATAATTGTGCTGTTTGTGGATGAAGACAGAAAGCAAGGAGCACCCCTGTGTACGAGC	1060
Qy	132	CATGGACAGCTGCTCTTAGAGATTGTAATGAGGAAACCCAGGAGACTCTCAATA	191
Db	1059	CATAGACAGCTGCTCTTGAAGTGGCTAACTGGAGGAAATACCAGGAGACTATCAATA	1000
Qy	192	CCAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAGACATCTATGAGGTTCACT	251
Db	999	TCAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGTGGGACATCTATGAGGTTTGT	940
Qy	252	GCAGCCCATCTTAAGGCACTTGTGGAAGGGCAGAAATGCCAGTGTGCTTATGGAC	311
Db	939	ACAGCCCATCTTAAGGCACTTGTGGAAGGGCAGAAATGCCAGTGTGCTTATGGAC	880
Qy	312	CACAGGAGCTGGGAGAGGAGCAGACATCTGTTGGGAGCCAGAGCACTGGGGTATCCC	371
Db	879	GACTGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	822
Qy	372	CGCGGCTCTCATCGACCTCTGAGCTCACAAGGAGGAGGAGGAGGAGGAGGAGGAGG	431
Db	821	TCAGACTCTCCAGAGCTCTGACCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGG	762
Qy	432	GGCCCTTTCTGTGTCACCAATGCTTACCTAGAGATCTACAGGAGAGGAGGAGGAGG	491
Db	761	GGACATTTCTGTTGCTTTGTGCTATTAGATCTCTACAGGAGGAGGAGGAGGAGG	702
Qy	492	GGACCTGCTTGGGAGAGCTGGTAATCCGAGAGAGTCCCGGGGAGGAGGAGGAGG	551
Db	701	GGATCTCTTATAGTGTACCTAAATGTCAGCGGAGGAGGAGGAGGAGGAGGAGG	642

Center project name: L18071
Center clone name: 202 A 19

FEATURES	Location/Qualifiers
source	1. .196674 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" /chromosome="1" /map="1" /clones="RP23-202A19" /clone_lib="RPCI-23 Female Mouse BAC"
misc_feature	1. .4229 /note="wgs end extension clone end:SP6"
repeat_region	1. .193
unseq	614. .746 /rpt_family="MTA" /note="single clone coverage"
unseq	701. .706 /note="30 qual single clone coverage"
repeat_region	1208. .1291 /rpt_family=" (CTA) n"
repeat_region	1571. .1842 /rpt_family="RLTRETN MM" complement(1974. .2067)
repeat_region	3219. .3319 /rpt_family="B4" /rpt_family="GA-rich"
repeat_region	3368. .3448 /rpt_family=" (CA) n"
repeat_region	3626. .3688 /rpt_family=" (CAAA) n"
repeat_region	3925. .4374 /rpt_family="Lx"
misc_feature	4230. .4235 /note="clone boundary clone end:SP6 site:ECORI"
repeat_region	4383. .4411 /rpt_family="AT-rich"
repeat_region	4445. .4496 /rpt_family=" (TCTCTG) n"
repeat_region	4509. .4638 /rpt_family=" (TG) n"
repeat_region	4741. .4784 /rpt_family=" (TG) n"
repeat_region	4785. .4832 /rpt_family=" (TA) n"
repeat_region	4833. .4955 /rpt_family=" (TG) n"
repeat_region	5398. .5431 /rpt_family=" (GAAA) n" complement(5434. .5561)
repeat_region	5561. .5611 /rpt_family="RMER15" complement(6185. .6272)
repeat_region	6516. .6570 /rpt_family="RSINE1" /rpt_family=" (T) n"
repeat_region	6571. .6823 /rpt_family="L1_MM" complement(6821. .6939)
repeat_region	7007. .7285 /rpt_family="L1_MM"
repeat_region	7292. .7848 /rpt_family="Lx5"
repeat_region	8187. .8286 /rpt_family="RSINE1" complement(8283. .9061)
repeat_region	9061. .9815 /rpt_family="B4A" complement(9243. .9815)

[illegible]

Fri Nov 12 12:26:24 2004

Qy	665	GTATGCTGTGCTCTCTGCTCAAGGTGGAACAGCGGGAAGTTTGGCCCCCATTTTGGCCAGC	724
Db	102231	-TCATTCATGCTCTTTATCAAGGTGGATCAGCATGGACATTTGACTCTGTTTCACTAGT	102289
Qy	725	GAGAGGGAACCTCTACCTGATTTGCTGGCTGAGGACAAACCGGCGCACAGCA	784
Db	102290	GAAAAGGAACCTCTACCATTTGATTTGGTTGGCTCAGAGGAAACCAACA-AGGA	102348
Qy	785	ACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGCA	844
Db	102349	ACAGTGCATTTCACTCAAGAGAGTGGGCCATCAACACCTGCTTC-TTGTACTGGCA	102407
Qy	845	AAGTGTAGATGGCTGAATCAGGGCTCCCTGCTGATCTTATCGGACAGCAAGCTCA	904
Db	102408	AAGTGGAGTTGCATCATATTAGGGCTCCCTCTATATATATCTGGGACAGAAAAATCA	102467
Qy	905	CTCGCTATTCAGGACTCTCTGGTGGCTCAGCCACAGTATCCTTATTTGCCAACATTG	964
Db	102468	TTTGGCTATTGCCAGACTCTCTGGAGGGCTCAGCTCATTAACATCATCATTTACCAACTG	102527
Qy	965	CCCTCAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGTCTGCCAGTCCA	1024
Db	102528	CTCTGAAAGACATTTTACTTGGATCATTTCTGACATTTGAATTTCACTGTAGTCCA	102587
Qy	1025	AGAGGTGATCAATCGGCTTTTACCAATGAGAGCTCGAGCTCATGCTTTGGGACCTG	1084
Db	102588	AGGAGCTGATTAATTTGGCTTTTCAACCAATGGAGTCTACAGCTCATGCTTACACACTG	102647
Qy	1085	TTAAGCTGCTCAGAAAGATTCCTTTGCTCCACGAGCAAGAGAGCCCGAGGCCCTG	1144
Db	102648	TTAAGCTTTGGCAAAAGAACTTCTAGGCCATCAAGGCAAGAAAGCAAGGTCTTG	102707
Qy	1145	AGGAAGAGAGATTGGGAGCCCTGAGCCCATGGAGCTCCAGCTCTGCTCCGAAAC	1204
Db	102708	TAGAAGATCAATGGGGAGTCTTAACCTACAGCAGTTTTATCAACTGGCTCCAGAAAC	102767
Qy	1205	TCAGCCCCCTACAGAGCTTAAGCAGCATGGACCGGCCATGCTGGAGCGCTCTCAGCT	1264
Db	102768	TAAGCTCTTTATAGAACCTAAGCAATGTGGATCCAGCTATGTTGGAGAGTTCCTGAGTT	102827
Qy	1265	TGGACCGCTGCTTGCTCCAGGGAGCCAG--GGGGCCCTCTGTGAGTACCCCAA	1322
Db	102828	TATATCAATTTGTTGGGATCTCAGGGGACAAAGAGGTATATTCTACTGAATACCCCAA	102887
Qy	1323	GCGAGA-GCGGATGCTTAATGAAGACAGTGAAGAGAGGACCTAGAGATTGAGAGC	1381
Db	102888	GCAAGACAAGATGGTACTCATTAAGACAGT-GTAGGAAAGAACTTGGAGATTGAGAGC	102946
Qy	1382	TTAAGACGAACCAAAAGAACTGGGGCCCAAGATTTGGCCCAAGGCTGAGGAAAAGG	1441
Db	102947	TTAAGGTGAAGTAGAAAGAACTATAGGCAAGGTGTTGGCACAGGATCCTAGATATGA	103006
Qy	1442	AG 1443	
Db	103007	AG 103008	

Search completed: November 10, 2004, 13:18:11
Job time : 6475.64 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2004, 09:45:10 ; Search time 756.977 Seconds
(without alignments)
10152.424 Million cell updates/sec

Title: US-10-797-893-7

Perfect score: 1464

Sequence: 1 atgggtcgctgtcggttaag.....accattgtccacaatgtga 1464

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1464	100.0	1464	10	ADC23343
2	1464	100.0	1464	12	ADQ60233
3	1461	99.8	1538	10	ADC23339
4	1461	99.8	1538	12	ADQ60229
5	1458	99.6	2097	12	ADQ09241
6	1447	98.8	2099	10	ADK40995
7	1443.2	98.6	2104	3	AAL15853
8	1038.4	70.9	1041	10	ADC23341
9	1038.4	70.9	1041	12	ADQ60231
10	1035.4	70.7	1115	10	ADC23337
11	1035.4	70.7	1115	12	ADQ60227
12	480	32.8	491	6	ABK70260
13	376.4	25.7	386	4	AAS38879
14	265	18.1	464	9	ACH44493
15	233.2	15.9	531	12	ACH77408
16	230.4	15.7	232	12	ACH91108
17	185.6	12.7	2095	11	ADM01564
18	185.6	12.7	3624	12	ADQ44181
19	156.8	10.7	3374	10	ADJ95073
20	154.8	10.6	2375	8	ABX34596
21	138.6	9.5	4108	6	ABA94614

22	138.6	9.5	4108	8	ABS57215	Abx57215 Human kin
23	137.2	9.4	2675	10	ADJ95077	Adj95077 Novel NOV
24	136.6	9.3	1152	6	ABA94616	Abx94616 Nucleotid
25	136.6	9.3	1152	8	ABS57216	Abx57216 DNA encod
26	128.8	8.8	897	4	AAI70088	Aai70088 Human kin
27	128.8	8.8	897	6	AAI39619	Aai39619 DNA encod
28	128.8	8.8	897	12	ADM81051	Adm81051 Human Hsk
29	127.8	8.7	488	5	ABV47991	Abv47991 Human pro
30	127.6	8.7	482	10	ABT40796	Abt40796 Toxicity
31	126.6	8.6	1026	6	ABQ73061	Abq73061 Human kin
32	126.6	8.6	1026	9	AAL56806	Aal56806 DNA encod
33	126.6	8.6	1026	10	ADB66785	Adb66785 DNA encod
34	124.2	8.5	1011	6	ABQ73062	Abq73062 Human kin
35	124.2	8.5	1011	9	AAL56807	Aal56807 DNA encod
36	124.2	8.5	1011	10	ADB66787	Adb66787 DNA encod
37	121.8	8.3	3661	6	AAD28568	Aad28568 Human kin
38	121.8	8.3	3694	10	ADC10189	Adc10189 Human NOV
39	121.4	8.3	1014	6	ABA94615	Abx94615 Human Hsk
40	121.2	8.3	3570	12	ADQ19012	Adq19012 Human sof
41	119.8	8.2	2994	4	ABL28591	Abi28591 Drosophil
42	119.4	8.2	2034	4	ABL16733	Abi16733 Drosophil
43	119.4	8.2	4034	4	ABL16732	Abi16732 Drosophil
44	119.2	8.1	2132	10	ADC30338	Adc30338 Human nov
45	117	8.0	1839	10	ADC30569	Adc30569 Human nov

ALIGNMENTS

RESULT 1

ADC23343
ID ADC23343 standard; DNA; 1464 BP.
XX
AC ADC23343;
XX
DT 18-DEC-2003 (first entry)
XX
DE DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 7).
XX
KW human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;
KW cytosolic; cardiant; immunomodulator; antiinflammatory; gene therapy;
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;
KW cardiac hypertrophy; immune disorder; inflammation.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1464
FT /tag= a
FT /product= "KID protein"
XX
DN US6387644-B1.

XX
PD 14-MAY-2002.
XX
PF 28-NOV-2000; 2000US-00724224.
XX
PR 20-APR-1999; 99US-00295612.
XX
PR 20-JUN-2000; 2000US-00597292.
XX
XX (CVTO-) CYTOKINETICS INC.
XX
PI Beraud C;
XX
XX WPI; 2003-706919/67.
XX
DR P-PSDB; ADC23344.
XX
PT Identifying a candidate agent as modulator of function of a target
PT protein for treating cellular proliferation disorders by adding a
PT candidate agent to a mixture of the target protein that
PT directly/indirectly produces ADP or phosphate.
XX
PS Disclosure; SEQ ID NO 7; 26pp; English.

XX This invention relates to a novel method for high throughput screening
 CC systems used to identify compounds for the treatment of cellular
 CC proliferation disorders. Specifically, it refers to candidate agents that
 CC are capable of modulating the activity of target proteins having motor
 CC domains, such that the target protein directly or indirectly produces ADP
 CC or phosphate. Furthermore, this activity can be determined using
 CC fluorescence or absorbance readouts. The present invention describes a
 CC method that identifies modulators of the target protein, which is a
 CC kinesin-like DNA binding protein (known as KID) as cytosolic, cardiac,
 CC immunomodulators and antiinflammatory. Accordingly, through gene
 CC therapy, they can be used for the treatment of cancer, hyperplasias,
 CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This
 CC polynucleotide sequence is human KID DNA (Seqid 7) encoding a full length
 CC KID enzyme of the invention.

SQ Sequence 1464 BP; 342 A; 409 C; 428 G; 285 T; 0 U; 0 Other;

Query Match 100.0%; Score 1464; DB 10; Length 1464;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1464; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGTCGCTGTCGGCTAAGCAAGATTGGAGTACTCGTCGTCACCTCCAGCTCGCGTA 60
 DB 1 ATGGGTCGCTGTCGGCTAAGCAAGATTGGAGTACTCGTCGTCACCTCCAGCTCGCGTA 60
 QY 61 AGGGTGGCTGTGCGAATGCGGCAATTGTGGATGGAAACAGCGGGAGCAAGTATCCCCC 120
 DB 61 AGGGTGGCTGTGCGAATGCGGCAATTGTGGATGGAAACAGCGGGAGCAAGTATCCCCC 120
 QY 121 TGTGTGGGGGATGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACACACAGGAG 180
 DB 121 TGTGTGGGGGATGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACACACAGGAG 180
 QY 181 ACTCTCAAAATACAGATTGTATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTAT 240
 DB 181 ACTCTCAAAATACAGATTGTATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTAT 240
 QY 241 GCAGGTTCACTGAGCCCATCTTAAGCACTTCTGTAAGGGGAGAGTCCAGTGTGCTT 300
 DB 241 GCAGGTTCACTGAGCCCATCTTAAGCACTTCTGTAAGGGGAGAGTCCAGTGTGCTT 300
 QY 301 GCCTATGGAACCCAGAGCTGGGAACGACACACATGCTGGGAGCCGCCAGCAACCT 360
 DB 301 GCCTATGGAACCCAGAGCTGGGAACGACACACATGCTGGGAGCCGCCAGCAACCT 360
 QY 361 GGGGTGATCCCGGGCTCTCATGGACCTCTCATGGACCTCTCATGGACCTCTCATGGAC 420
 DB 361 GGGGTGATCCCGGGCTCTCATGGACCTCTCATGGACCTCTCATGGACCTCTCATGGAC 420
 QY 421 GGCGGGCATGGGGCCCTTCTGTACCATGCTTACCTAGAGATCTACAGGAGAGGTA 480
 DB 421 GGCGGGCATGGGGCCCTTCTGTACCATGCTTACCTAGAGATCTACAGGAGAGGTA 480
 QY 481 TTAGACCTCTCGGACCTGCTTCCGAGACCTGCTTAATCCGAGAACTGCGCGGGGAAT 540
 DB 481 TTAGACCTCTCGGACCTGCTTCCGAGACCTGCTTAATCCGAGAACTGCGCGGGGAAT 540
 QY 541 ATCTGATTCGGGTCTCTCCAGAACCCATCATAGTACTTGTGATTTTGGAGCGGCAC 600
 DB 541 ATCTGATTCGGGTCTCTCCAGAACCCATCATAGTACTTGTGATTTTGGAGCGGCAC 600
 QY 601 TTCTGTCCAGCAGCTCGAAATCGACTGTAGGACCCACCGGCTCAACAGCGCTCTCTCC 660
 DB 601 TTCTGTCCAGCAGCTCGAAATCGACTGTAGGACCCACCGGCTCAACAGCGCTCTCTCC 660
 QY 661 CGCAGTCACTGTGCTCTCTGCTCAAGGTGGACAGCGGGAACTTTGGCCCCATTTCCG 720
 DB 661 CGCAGTCACTGTGCTCTCTGCTCAAGGTGGACAGCGGGAACTTTGGCCCCATTTCCG 720
 QY 721 CAGCGAGAGGAAATCTTACTGATTTGACTTGGCTCAGAGGACCAACCGCGGCACA 780
 DB 721 CAGCGAGAGGAAATCTTACTGATTTGACTTGGCTCAGAGGACCAACCGCGGCACA 780

QY 781 GGCAACAGGCGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTCTTTGTCCTG 840
 DB 781 GGCAACAGGCGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTCTTTGTCCTG 840
 QY 841 GGCAAGTGGTAGATGGCTGAATCAGGGCCCTCCCTGCTGTACCTTATCGGAGCAGCAAG 900
 DB 841 GGCAAGTGGTAGATGGCTGAATCAGGGCCCTCCCTGCTGTACCTTATCGGAGCAGCAAG 900
 QY 901 CTCACTCCCTTATTCAGGACTCTCTGGGTGGCTCAGGCCACAGTATCTTATTCGCAAC 960
 DB 901 CTCACTCCCTTATTCAGGACTCTCTGGGTGGCTCAGGCCACAGTATCTTATTCGCAAC 960
 QY 961 ATTGCCCCCTGAGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGG 1020
 DB 961 ATTGCCCCCTGAGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGG 1020
 QY 1021 TCCAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTCAGCCTTCATGCTTGGGA 1080
 DB 1021 TCCAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTCAGCCTTCATGCTTGGGA 1080
 QY 1081 CCTGTTAAGCTGCTCAGAAAGATTGGTCCACAGAGGCAAGAGAGCCGAGGC 1140
 DB 1081 CCTGTTAAGCTGCTCAGAAAGATTGGTCCACAGAGGCAAGAGAGCCGAGGC 1140
 QY 1141 CCTGAGGAGAGGAGATTGGAGCCCTCAGGCCATGAGCCTTCCAGCCTTCGCTCCCAG 1200
 DB 1141 CCTGAGGAGAGGAGATTGGAGCCCTCAGGCCATGAGCCTTCCAGCCTTCGCTCCCAG 1200
 QY 1201 AAATCAGCCCTTACAGAGCTAAGAGCTATGGACCCGAGCCTTCCAGCCTTCGCTCCCAG 1260
 DB 1201 AAATCAGCCCTTACAGAGCTAAGAGCTATGGACCCGAGCCTTCCAGCCTTCGCTCCCAG 1260
 QY 1261 AGCTTGGACCGCTGCTTCCCTCCAGGGAGCCAGGGGCCCTCTGTTGATGATCCCA 1320
 DB 1261 AGCTTGGACCGCTGCTTCCCTCCAGGGAGCCAGGGGCCCTCTGTTGATGATCCCA 1320
 QY 1321 AAGCGAGAGCGGATGCTGCTAATGAAGACAGTGAAGAGAGGACCTAGAGATTGAGAGG 1380
 DB 1321 AAGCGAGAGCGGATGCTGCTAATGAAGACAGTGAAGAGAGGACCTAGAGATTGAGAGG 1380
 QY 1381 CTTAAGACGAGCAAAAGAACTGAGGCCAGATGTTGGCCAGAGGCTCAGGAGAAAG 1440
 DB 1381 CTTAAGACGAGCAAAAGAACTGAGGCCAGATGTTGGCCAGAGGCTCAGGAGAAAG 1440
 QY 1441 GAGAACCTTGTCCCAATGTGA 1464
 DB 1441 GAGAACCTTGTCCCAATGTGA 1464

RESULT 2

ADQ60233
 ID ADQ60233 standard; DNA; 1464 BP.

XX ADQ60233;

XX AC

XX 23-SEP-2004 (first entry)

XX Human microtubule motor protein DNA #4.

XX Human; microtubule motor protein; gene; ds;
 KW cellular proliferation disorder; cancer; hyperplasia; restenosis;
 KW cardiac hypertrophy; immune disorder; inflammation;
 KW kinesin-like DNA binding protein; KID; autoimmune disease; arthritis;
 KW graft rejection; inflammatory bowel disease; angioplasty.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..1464

XX FT /*tag= a

XX FT /product= "Human microtubule motor protein #4"

XX XX

PN US6762043-B1.
 PD 13-JUL-2004.
 XX 06-MAR-2002; 2002US-00093317.
 XX 20-APR-1999; 99US-00295612.
 PR 20-JUN-2000; 2000US-00597292.
 PR 28-NOV-2000; 2000US-00724224.
 XX (CVTO-) CYTOKINETICS INC.
 PA Beraud C;
 XX WPI; 2004-532491/51.
 DR P-PSDB; ADQ60234.
 XX
 PT New isolated microtubule motor protein, useful for screening modulators
 PT for treating cellular proliferation disorders such as cancer,
 PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
 PT inflammation.
 XX
 PS Disclosure; SEQ ID NO 7; 26pp; English.
 XX
 CC The invention relates to human microtubule motor proteins and the nucleic
 CC acids encoding them. The invention also relates to a method of screening
 CC for modulators of a motor protein which has microtubule stimulated ATPase
 CC activity, a method of testing for ATPase activity of microtubule motor
 CC proteins, methods to identify candidate agents that bind to a target
 CC protein or act as a modulator of the binding characteristics or
 CC biological activity of a target protein, modulators of the target
 CC protein, and methods of treating cellular proliferation disorders such as
 CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
 CC and inflammation, for treating disorders associated with kinesin-like DNA
 CC binding protein (KID) and for inhibiting KID. The sequences are used for
 CC screening for modulators of motor proteins useful for treating cellular
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
 CC hypertrophy, immune disorders and inflammation, for treating disorders
 CC associated with KID and for inhibiting KID and for treating autoimmune
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and
 CC proliferation induced after medical procedures including surgery and
 CC angioplasty. This sequence represents DNA encoding a human microtubule
 CC motor protein of the invention.
 XX
 SQ Sequence 1464 BP; 342 A; 409 C; 428 G; 285 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1464; DB 12; Length 1464;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1464; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ATGGGTGCTGCTCGGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGGTA 60
 Db 1 ATGGGTGCTGCTCGGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGGTA 60
 Qy 61 AGGGTGTGCTGCACTCGGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGGTA 120
 Db 61 AGGGTGTGCTGCACTCGGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGGTA 120
 Qy 121 TGTGTGGGGCATGGACAGCTGCTCTCTAGAGATTGCTTAAGTGGAGAACACACAGAG 180
 Db 121 TGTGTGGGGCATGGACAGCTGCTCTCTAGAGATTGCTTAAGTGGAGAACACACAGAG 180
 Qy 181 ACTCTCAATACCAAGTTTGTATGCTTCTATGGGGAGAGGAGTACTCAGCAGGACATCTAT 240
 Db 181 ACTCTCAATACCAAGTTTGTATGCTTCTATGGGGAGAGGAGTACTCAGCAGGACATCTAT 240
 Qy 241 GCAGGTTCAAGTGGAGCCCATCTTAAGGACCTTGTCTGAAGGGCAGAAATGCCAGTGTCTT 300
 Db 241 GCAGGTTCAAGTGGAGCCCATCTTAAGGACCTTGTCTGAAGGGCAGAAATGCCAGTGTCTT 300
 Qy 301 GCCTATGGACCCACAGAGCTGGGAGACGACACAAATGCTGGGCGACCCAGAGCAACCT 360
 Db 301 GCCTATGGACCCACAGAGCTGGGAGACGACACAAATGCTGGGCGACCCAGAGCAACCT 360

Qy 361 GGGGTGATCCCGGGCTCTCATGGACCTCTCTGAGCTCAAAAGGAGAGGGTCCCGAG 420
 Db 361 GGGGTGATCCCGGGCTCTCATGGACCTCTCTGAGCTCAAAAGGAGAGGGTCCCGAG 420
 Qy 421 GGGCGGCATGGGCGCTTTCTGTCAACATGCTTACCTAGAGATCTACAGAGAGGTA 480
 Db 421 GGGCGGCATGGGCGCTTTCTGTCAACATGCTTACCTAGAGATCTACAGAGAGGTA 480
 Qy 481 TTAGACCTCTGAGACCTGCTTGGGAGACCTGCTTATCCGAGAGACTGCGGGGGAT 540
 Db 481 TTAGACCTCTGAGACCTGCTTGGGAGACCTGCTTATCCGAGAGACTGCGGGGGAT 540
 Qy 541 ATCTGTGATTCGGGCTCTCTCCAGAACCCATAGTAGCTTTGTGATTTTGTAGCGGCAC 600
 Db 541 ATCTGTGATTCGGGCTCTCTCCAGAACCCATAGTAGCTTTGTGATTTTGTAGCGGCAC 600
 Qy 601 TTCTGCGCAGCAGTCGAAATCGAATCGAGTGTAGAGCCACCCGGCTCAACAGCGTCTCC 660
 Db 601 TTCTGCGCAGCAGTCGAAATCGAATCGAATCGAGTGTAGAGCCACCCGGCTCAACAGCGTCTCC 660
 Qy 661 CGCAGTCATGCTGCTCTGCTCAAGTGTGACCGAGCGGGAACGTTTGGCCCCATTTCCG 720
 Db 661 CGCAGTCATGCTGCTCTGCTCAAGTGTGACCGAGCGGGAACGTTTGGCCCCATTTCCG 720
 Qy 721 CAGCGAGAGGGAACCTCTACCTGATTTGCTGCTGAGGACCAACCGGCGCACA 780
 Db 721 CAGCGAGAGGGAACCTCTACCTGATTTGCTGCTGAGGACCAACCGGCGCACA 780
 Qy 781 GGCNAACAGGCGCTTCGGCTAAAAGAGAGTGGAGCCATCAACACCTCTCTGTTGCTCG 840
 Db 781 GGCNAACAGGCGCTTCGGCTAAAAGAGAGTGGAGCCATCAACACCTCTCTGTTGCTCG 840
 Qy 841 GGCNAAGTGTAGTGGCTGATCAGGCGCTCTCTGTTGCTGATCTTATCGGAGCAGCAAG 900
 Db 841 GGCNAAGTGTAGTGGCTGATCAGGCGCTCTCTGTTGCTGATCTTATCGGAGCAGCAAG 900
 Qy 901 CTCACTCGCTTATTCAGGAGCTCTCTGGTGGCTCAGCCACAGTATCTTTATTTGCCAAC 960
 Db 901 CTCACTCGCTTATTCAGGAGCTCTCTGGTGGCTCAGCCACAGTATCTTTATTTGCCAAC 960
 Qy 961 ATTGCCCTCAGAGAGCGCTTCTACCTAGACACAGTCTCCGACCTCAACCTTTGCTGCCAGG 1020
 Db 961 ATTGCCCTCAGAGAGCGCTTCTACCTAGACACAGTCTCCGACCTCAACCTTTGCTGCCAGG 1020
 Qy 1021 TCCAGAGAGTGTATTCGGCTTTTACCAATGAGAGCGCTGAGCGCTCATGCTTGGGA 1080
 Db 1021 TCCAGAGAGTGTATTCGGCTTTTACCAATGAGAGCGCTGAGCGCTCATGCTTGGGA 1080
 Qy 1081 CCTGTTAAGCTGTCTCAGAAAGAAATTCCTTGGTCCACAGAGGCAAGAGAGCCGAGGC 1140
 Db 1081 CCTGTTAAGCTGTCTCAGAAAGAAATTCCTTGGTCCACAGAGGCAAGAGAGCCGAGGC 1140
 Qy 1141 CTGAGGAAGAGAGATTTGGAGCCCTGAGCCATGAGCTCCAGCTCTGCTTCCCGAG 1200
 Db 1141 CTGAGGAAGAGAGATTTGGAGCCCTGAGCCATGAGCTCCAGCTCTGCTTCCCGAG 1200
 Qy 1201 AAATCAGCCCCCTACAGAAAGTGAAGCAGCATGAGCCGGGCGATGCTGGAGCGCTCTC 1260
 Db 1201 AAATCAGCCCCCTACAGAAAGTGAAGCAGCATGAGCCGGGCGATGCTGGAGCGCTCTC 1260
 Qy 1261 AGCTTGGAGCGCTCTGCTTGGCTCCCGGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCA 1320
 Db 1261 AGCTTGGAGCGCTCTGCTTGGCTCCCGGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCA 1320
 Qy 1321 AAGCGAGAGCGGATGCTGCTAATGAAGACAGTGAAGAGAGGACCTTAGAGATTGAGAGG 1380
 Db 1321 AAGCGAGAGCGGATGCTGCTAATGAAGACAGTGAAGAGAGGACCTTAGAGATTGAGAGG 1380
 Qy 1381 CTTAAGACGAGCAAAAGAACTGGAGGCGCAAGATGTTGGCCCAAGAGGCTGAGGAAAG 1440
 Db 1381 CTTAAGACGAGCAAAAGAACTGGAGGCGCAAGATGTTGGCCCAAGAGGCTGAGGAAAG 1440

QY	1441	GAGAACCAATTGTCACCAATGTA	1464	QY	4	GGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTGCTCCACCTCCAGCTCGCGTAAGG	63
DB	1441	GAGAACCAATTGTCACCAATGTA	1464	DB	78	GGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTGCTCCACCTCCAGCTCGCGTAAGG	137
RESULT 3				QY	64	GTGGCTGTGGCACTGCGGCCAATTTGTGGATGGAAACAGCGGAGCAAGTGAATCCCCCTGT	123
ADC23339				DB	138	GTGGCTGTGGCACTGCGGCCAATTTGTGGATGGAAACAGCGGAGCAAGTGAATCCCCCTGT	197
XX				QY	124	GTGGGGGCATGGACAGCTGCTCTAGAGAGTTCTTAACCTGGAGGAAACCAAGAGACT	183
AC				DB	198	GTGGGGGCATGGACAGCTGCTCTAGAGAGTTCTTAACCTGGAGGAAACCAAGAGACT	257
XX	18-DEC-2003	(first entry)		QY	184	CTCAATAACCAAGTTTGTATGCTCTTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	243
DE				DB	258	CTCAATAACCAAGTTTGTATGCTCTTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	317
XX				QY	244	GGTTCAAGTGCAGCCCATCTTAAGGCACTTGTGGAAGGCGAGAAATGCCAGTGTGCTTCCC	303
KW				DB	318	GGTTCAAGTGCAGCCCATCTTAAGGCACTTGTGGAAGGCGAGAAATGCCAGTGTGCTTCCC	377
KW				QY	304	TATGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGCCACAGAGCAACTCGG	363
OS				DB	378	TATGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGCCACAGAGCAACTCGG	437
XX				QY	364	GTGATCCCGCGGCTCTCATGGACCTCTCTGAGCTCACAAGGAGGAGGGTGCAGGGC	423
PH				DB	438	GTGATCCCGCGGCTCTCATGGACCTCTCTGAGCTCACAAGGAGGAGGGTGCAGGGC	497
FT				QY	424	CGGCCATGGGCGCTTCTGTCAACATGCTTACCTAGAGATCTACAGAGAGAGTATTA	483
FT				DB	498	CGGCCATGGGCGCTTCTGTCAACATGCTTACCTAGAGATCTACAGAGAGAGTATTA	557
XX	US6387644-B1.			QY	484	GACCTCTGGACCCCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTGCCGGGGGAATATC	543
XX	14-MAY-2002.			DB	558	GACCTCTGGACCCCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTGCCGGGGGAATATC	617
XX	28-NOV-2000; 2000US-00724224.			QY	544	CTGATTCGGGCTCTCTCCAGAAAGCCATAGTAGCTTGTGTAATTTGAGGGGCACTTC	603
XX	20-APR-1999; 99US-00295612.			DB	618	CTGATTCGGGCTCTCTCCAGAAAGCCATAGTAGCTTGTGTAATTTGAGGGGCACTTC	677
PR	20-JUN-2000; 2000US-00597292.			QY	604	CTGCCAGCCAGTCGAATTCGACTAGGAGCCACCGCTCAACAGCGCTCTCCCGC	663
XX	(CYTO-) CYTOKINETICS INC.			DB	678	CTGCCAGCCAGTCGAATTCGACTAGGAGCCACCGCTCAACAGCGCTCTCCCGC	737
XX	Beraud C;			QY	664	AGTCATGCTGCTCTCTGCTCAAGTGGACAGCGGAAAGCTTTGGCCCCCATTTGCCAG	723
XX	WPI; 2003-706919/67.			DB	738	AGTCATGCTGCTCTCTGCTCAAGTGGACAGCGGAAAGCTTTGGCCCCCATTTGCCAG	797
DR	P-PSDB; ADC23340.			QY	724	CGAGAGGGGAAAACTCTACCTGATTGACTTGGCTGGGTCAAGAGCAACCCGGCGCACAGC	783
XX	Identifying a candidate agent as modulator of function of a target			DB	798	CGAGAGGGGAAAACTCTACCTGATTGACTTGGCTGGGTCAAGAGCAACCCGGCGCACAGC	857
PT	protein for treating cellular proliferation disorders by adding a			QY	784	AACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCCCTGTTTGTCTCGGC	843
PT	candidate agent to a mixture of the target protein that			DB	858	AACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCCCTGTTTGTCTCGGC	917
PT	directly/indirectly produces ADP or phosphate.			QY	844	AAAGTGTAGATGCGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGAGCAAGAGCTC	903
XX	Disclosure; SEQ ID NO 3; 26pp; English.			DB	918	AAAGTGTAGATGCGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGAGCAAGAGCTC	977
XX	This invention relates to a novel method for high throughput screening			QY	904	ACTCGCTATTGTCAGAGCTCTCTGGTGGCTGAGCCACAGTATCTTATTCGCAACATT	963
XX	systems used to identify compounds for the treatment of cellular			DB	978	ACTCGCTATTGTCAGAGCTCTCTGGTGGCTGAGCCACAGTATCTTATTCGCAACATT	1037
CC	proliferation disorders. Specifically, it refers to candidate agents that			QY	964	GCCCTGTAGAGAGCTTCTTACCTAGACAGCTCTCCGCACTCAACTTTTGTGCGAGGTCC	1023
CC	are capable of modulating the activity of target proteins having motor			DB	1038	GCCCTGTAGAGAGCTTCTTACCTAGACAGCTCTCCGCACTCAACTTTTGTGCGAGGTCC	1097
CC	domains, such that the target protein directly or indirectly produces ADP			QY	1024	AAGGAGGTGATCAATCGGCTTTTATCAATGAGAGCTGACGCTCATGCTTGGGACCT	1083
CC	or phosphate. Furthermore, this activity can be determined using			DB	1098	AAGGAGGTGATCAATCGGCTTTTATCAATGAGAGCTGACGCTCATGCTTGGGACCT	1157
CC	fluorescence or absorbance readouts. The present invention describes a			QY			
CC	method that identifies modulators of the target protein, which is a			DB			
CC	kinesin-like DNA binding protein (known as KID) as cytoskeletal, cardiac,			QY			
CC	immunomodulators and antiinflammatories. Accordingly, through gene			DB			
CC	therapy, they can be used for the treatment of cancer, hyperplasias,			QY			
CC	retenosis, cardiac hypertrophy, immune disorders and inflammation. This			DB			
CC	polynucleotide sequence is human KID DNA (SeqID 3) encoding a full length			QY			
CC	KID enzyme of the invention.			DB			
XX	Sequence 1538 BP; 355 A; 432 C; 459 G; 292 T; 0 U; 0 Other;			QY			
XX	Query Match 99.8%; Score 1461; DB 10; Length 1538;			DB			
XX	Best Local Similarity 100.0%; Pred. No. 0;			DB			
XX	Matches 1461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			DB			

QY 724 CGAGAGGGAAGAACTCTACCTGATTGACTTGGCTGGGTGAGGACAAACGGGGCCACAGGC 783
DB 798 CGAGAGGGAAGAACTCTACCTGATTGACTTGGCTGGGTGAGGACAAACGGGGCCACAGGC 857
QY 784 AACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTCTGTTGTCTCTGGGC 843
DB 858 AACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTCTGTTGTCTCTGGGC 917
QY 844 AAGTGGTATAGTCGCTGAATCAGGGCCTCCCTCGTGTACCTTATCGGACACGAAGCTC 903
DB 918 AAGTGGTATAGTCGCTGAATCAGGGCCTCCCTCGTGTACCTTATCGGACACGAAGCTC 977
QY 904 ACTCGCTATTGAGAGACTCTCTGGGTGGCTCAGCCCAAGTATCTTATTGGCAACATT 963
DB 978 ACTCGCTATTGAGAGACTCTCTGGGTGGCTCAGCCCAAGTATCTTATTGGCAACATT 1037
QY 964 GCCCTTGAGAGAGCTTCTACCTAGACACAGTCTCGCACTCAACTTGTCTCCAGGCTC 1023
DB 1038 GCCCTTGAGAGAGCTTCTACCTAGACACAGTCTCGCACTCAACTTGTCTCCAGGCTC 1097
QY 1024 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTCGAGCTCATGCTTTGGACCT 1083
DB 1098 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTCGAGCTCATGCTTTGGACCT 1157
QY 1084 GTTAAGCTCTCAGAAAGAAATGCTTGTGTCACAGAGGCAAGAGAGCCCGAGGCCCT 1143
DB 1158 GTTAAGCTCTCAGAAAGAAATGCTTGTGTCACAGAGGCAAGAGAGCCCGAGGCCCT 1217
QY 1144 GAGGAGAGGAGATTGGGAGCCCTGAGCCCATGGCAGCTCCAGCTCTCTCCCTCCAGAAA 1203
DB 1218 GAGGAGAGGAGATTGGGAGCCCTGAGCCCATGGCAGCTCCAGCTCTCTCCCTCCAGAAA 1277
QY 1204 CTCAGCCCCCTCAGAAAGCTTAAGCAGCATGAGCCCGCCCATGCTGAGGCCCTCTCTCAGC 1263
DB 1278 CTCAGCCCCCTCAGAAAGCTTAAGCAGCATGAGCCCGCCCATGCTGAGGCCCTCTCTCAGC 1337
QY 1264 TTGAGCGCTGCTGCTCCCTCCAGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1323
DB 1338 TTGAGCGCTGCTGCTCCCTCCAGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1397
QY 1324 CGAGAGCGGATGCTGCTTAAGACAGTAGAGAGAGGACCTAGAGATTGAGAGGCTT 1383
DB 1398 CGAGAGCGGATGCTGCTTAAGACAGTAGAGAGAGGACCTAGAGATTGAGAGGCTT 1457
QY 1384 AAGACGAAGCAAAAGAACTGGAGGCCCAAGATTTGGCCCAAGAGCTGAGGAAAGGAG 1443
DB 1458 AAGACGAAGCAAAAGAACTGGAGGCCCAAGATTTGGCCCAAGAGCTGAGGAAAGGAG 1517
QY 1444 AACCATTTCCCAATGTGA 1464
DB 1518 AACCATTTCCCAATGTGA 1538

RESULT 5
ADQ09241
ID ADQ09241 standard; cDNA; 2097 BP.
XX AC ADQ09241;
XX AC ADQ09241;
DT 23-SEP-2004 (first entry)
XX Human KNSL4 encoding cDNA SEQ ID NO:426.
XX thanatos-associated protein; THAP; THAP responsive gene; THAP family;
KW THAP responsive element; angiogenesis; inflammation; metastasis; cancer;
KW apoptosis; cardiovascular disease; neurodegenerative disease; chemokine;
KW angiogenic; antiinflammatory; cardiovascular; cytostatic;
KW neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;
KW human; gene; ss.
XX Homo sapiens.
OS
XX W02004055050-A2.
PN

XX 01-JUL-2004.
XX PD 10-DEC-2003; 2003WO-IB006434.
XX PF 10-DEC-2002; 2002US-0432699P.
XX PR 03-JUL-2003; 2003US-0485027P.
XX PR (ENDO-) ENDOCUBE SAS.
XX PA (CNRS) CNRS CENT NAT RECH SCI.
XX PI Girard J, Amalric F, Roussigne M, Clouaire T;
XX WPI; 2004-525034/50.
XX DR P-PSDB; ADQ09240.
XX
PT Modulating expression of a Thanatos (death)-Associated Protein (THAP)
PT responsive gene for preventing or treating e.g. cancer or inflammation,
PT comprises modulating the interaction of a THAP polypeptide with a nucleic
PT acid.
XX Example 47; SEQ ID NO 426; 612pp; English.
XX
CC The present invention describes a method for modulating the expression of
CC a thanatos (death)-associated protein (THAP) responsive gene. The method
CC comprises modulating the interaction of a THAP-family polypeptide or its
CC biological fragment with a nucleic acid, and so enhancing or repressing
CC the expression of the THAP responsive gene. Also described: (1) a method
CC of modulating the expression of a gene responsive to a THAP/chemokine
CC complex; (2) a pharmaceutical composition comprising a THAP responsive
CC element in a pharmaceutical carrier; (3) a transcription factor decoy
CC consisting essentially of a THAP responsive element; (4) a cell
CC comprising a transcription factor decoy described above; (5) methods of
CC modulating the interaction between a nucleic acid and a THAP-family
CC polypeptide or its biological fragment, or a nucleic acid and a
CC THAP/chemokine complex; (6) a vector packaging cell line comprising a
CC cell comprising a viral vector which comprises a promoter operably linked
CC to a nucleic acid encoding a THAP-family polypeptide or its biological
CC fragment; (7) a method of constructing a cell which expresses a
CC recombinant THAP-family polypeptide; (8) a method of ameliorating
CC symptoms associated with a condition mediated by a THAP/chemokine complex
CC; (9) methods of identifying a test compound that modulates transcription
CC at a THAP responsive element or that modulates the transport of a
CC chemokine into the nucleus; (10) methods for reducing the symptoms
CC associated with a condition selected from excessive or insufficient
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
CC or insufficient apoptosis, cardiovascular disease and neurodegenerative
CC diseases; symptoms associated with a condition resulting from the
CC activity of a chemokine or a THAP-family polypeptide in an individual; or
CC symptoms associated with transcriptional repression or activation
CC mediated by a THAP-family polypeptide in an individual; (11) a vector
CC comprising a THAP responsive promoter operably linked to a nucleic acid
CC encoding a detectable product; (12) a genetically engineered cell
CC comprising the vector described above or that expresses a THAP-family
CC polypeptide or its biological fragment; (13) an in vitro transcription
CC reaction comprising a nucleic acid comprising a THAP responsive promoter,
CC ribonucleotides and an RNA polymerase; and (14) an isolated mutant THAP-
CC family polypeptide that does not bind to a chemokine. The pharmaceutical
CC composition has antiangiogenic, antiinflammatory, cardiovascular,
CC cytoskeletal, neuroprotective and osteopathic activities, and can be used
CC as a THAP and THAP synthesis modulator. The composition can be used for
CC modulating the expression of a THAP responsive gene. Modulation is useful
CC for reducing symptoms of conditions such as excessive or insufficient
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
CC or insufficient apoptosis, cardiovascular disease or neurodegenerative
CC diseases. The present sequence is used in the exemplification of the
CC present invention.
XX Sequence 2097 BP; 483 A; 592 C; 628 G; 394 T; 0 U; 0 Other;
SQ

Query Match 99.6%; Score 1458; DB 12; Length 2097;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GGTGCTGCTCGGCTAAGCAAGATTGGAGTACTCGTGTCCACCTCCAGCTCGGTAAGG 63
 Db 100 GGTGCTGCTCGGCTAAGCAAGATTGGAGTACTCGTGTCCACCTCCAGCTCGGTAAGG 159
 Qy 64 GTGGCTGTGCGACTCGGCTAATTTGTGATGCAAGAGGAGCAAGTATCCCTCTGT 123
 Db 160 GTGGCTGTGCGACTCGGCTAATTTGTGATGCAAGAGGAGCAAGTATCCCTCTGT 219
 Qy 124 GTGGGGGATGAGCAGCTGCTCTAGAGATTGCTAACTGAGAGAACCCACAGAGACT 183
 Db 220 GTGGGGGATGAGCAGCTGCTCTAGAGATTGCTAACTGAGAGAACCCACAGAGACT 279
 Qy 184 CTCAAATACCAATTTGATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 243
 Db 280 CTCAAATACCAATTTGATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 339
 Qy 244 GGTTCAGTGCAGCCCATCTTAAGCACTTGTGGAAGGAGAGTATGCAAGTGTCTTGGC 303
 Db 340 GGTTCAGTGCAGCCCATCTTAAGCACTTGTGGAAGGAGAGTATGCAAGTGTCTTGGC 399
 Qy 304 TATGAGCCCAAGAGCTGGGAAGACGACACATGCTGGGAGGAGGAGTGTGGG 363
 Db 400 TATGAGCCCAAGAGCTGGGAAGACGACACATGCTGGGAGGAGGAGTGTGGG 459
 Qy 364 GTGATCCCGGGCTCTCATGAGACTCTGAGCTCTCAAGAGGAGGAGTGTGGG 423
 Db 460 GTGATCCCGGGCTCTCATGAGACTCTGAGCTCTCAAGAGGAGGAGTGTGGG 519
 Qy 424 CGGCATGGGCTCTTCTGACCATCTTACCATGCTTACCTAGAGTCTACAGGAGAGTATTA 483
 Db 520 CGGCATGGGCTCTTCTGACCATCTTACCATGCTTACCTAGAGTCTACAGGAGAGTATTA 579
 Qy 484 GACCTCTGAGCCCTGCTTGGGAGACTTGTATATCCGAGAGACTCGCGGGGAATATC 543
 Db 580 GACCTCTGAGCCCTGCTTGGGAGACTTGTATATCCGAGAGACTCGCGGGGAATATC 639
 Qy 544 CTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTGAGCGGCACTTC 603
 Db 640 CTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTGAGCGGCACTTC 699
 Qy 604 CTGCGCCAGTTCGAATTCGATGTAGAGCCACCCGGTCAACAGAGCTCTCTCCGC 663
 Db 700 CTGCGCCAGTTCGAATTCGATGTAGAGCCACCCGGTCAACAGAGCTCTCTCCGC 759
 Qy 664 AGTCATCTGCTCTGCTCAAGTGCAGCGGAGCGTTTGGCCCATTTCCGAG 723
 Db 760 AGTCATCTGCTCTGCTCAAGTGCAGCGGAGCGTTTGGCCCATTTCCGAG 819
 Qy 724 CGAGAGGGAATCTCTACCTGATTGACTTGGCTGGGTGAGAGGACACCGGCGCAGGC 783
 Db 820 CGAGAGGGAATCTCTACCTGATTGACTTGGCTGGGTGAGAGGACACCGGCGCAGGC 879
 Qy 784 AACAGGGCTCTGGCTTAAAGAGAGTGGAGCCATCAACCTCTGTTGCTGGGC 843
 Db 880 AACAGGGCTCTGGCTTAAAGAGAGTGGAGCCATCAACCTCTGTTGCTGGGC 939
 Qy 844 AAGTGTGATGCTGCTGATCAGGGCTCTCTGCTGCTTACCTTACGGGACGACGAGCTC 903
 Db 940 AAGTGTGATGCTGCTGATCAGGGCTCTCTGCTGCTTACCTTACGGGACGACGAGCTC 999
 Qy 904 ACTCGCTATTGAGGAGCTCTCTGGGTGGCTCAGGCCACAGTATCTCTTATGCGCAACT 963
 Db 1000 ACTCGCTATTGAGGAGCTCTCTGGGTGGCTCAGGCCACAGTATCTCTTATGCGCAACT 1059
 Qy 964 GCCCTGTAGAGAGCTCTCTACCTAGACACAGTCTCCGCACTCAACTTTGTCGAGGTC 1023
 Db 1060 GCCCTGTAGAGAGCTCTCTACCTAGACACAGTCTCCGCACTCAACTTTGTCGAGGTC 1119
 Qy 1024 AAGAGGTGATCAATCGGCTTTTACCAATGAGGCTGAGCTCATGCTTGGGACCT 1083
 Db 1120 AAGAGGTGATCAATCGGCTTTTACCAATGAGGCTGAGCTCATGCTTGGGACCT 1179

Qy 1084 GTTAAAGCTGTCTCAGAAAGAAATTCCTTGGTCCACAGAGGCAAGAGAGCCCGAGGCCCT 1143
 Db 1180 GTTAAAGCTGTCTCAGAAAGAAATTCCTTGGTCCACAGAGGCAAGAGAGCCCGAGGCCCT 1239
 Qy 1144 GAGGAAGAGAGATTTGGAGCCCTTGAGCCCATGAGCTCCAGGCTCTGCTTCCAGAAA 1203
 Db 1240 GAGGAAGAGAGATTTGGAGCCCTTGAGCCCATGAGCTCCAGGCTCTGCTTCCAGAAA 1299
 Qy 1204 CTCAGCCCTCTACAGAAAGCTAAGCAGCATGAGCCCGGCTGCTGGAGCGCTCTCTCAGC 1263
 Db 1300 CTCAGCCCTCTACAGAAAGCTAAGCAGCATGAGCCCGGCTGCTGGAGCGCTCTCTCAGC 1359
 Qy 1264 TTGACCGTCTGCTTGGCTCTCCAGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1323
 Db 1360 TTGACCGTCTGCTTGGCTCTCCAGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1419
 Qy 1324 CGAGAGCGGATGCTGCTTAATGAGACAGTGAAGAGAGGACCTTAGAGATTGAGAGGCTT 1383
 Db 1420 CGAGAGCGGATGCTGCTTAATGAGACAGTGAAGAGAGGACCTTAGAGATTGAGAGGCTT 1479
 Qy 1384 AAGACGAAGCAAAAGAACTGGAGGCCAAGATGTTGGCCCAAGAGGCTGAGGAAAGGAG 1443
 Db 1480 AAGACGAAGCAAAAGAACTGGAGGCCAAGATGTTGGCCCAAGAGGCTGAGGAAAGGAG 1539
 Qy 1444 AACATTGTCCCAATG 1461
 Db 1540 AACATTGTCCCAATG 1557

RESULT 6
 ADK40995
 ID ADK40995 standard; DNA; 2099 BP.
 XX
 AC ADK40995;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Novel human kinase gene #15.
 XX
 KW cytosolic; immunomodulator; cardiant; neuroprotective; nootropic;
 KW analgesic; hypotensive; immunosuppressive; kinase inhibitor; kinase;
 KW cancer; peripheral nervous system; central nervous system;
 KW Alzheimer's disease; Parkinson's disease; multiple sclerosis;
 KW anyotrophic lateral sclerosis; viral infection; prion infection;
 KW ocular disease; migraine; pain; sexual dysfunction; mood disorder;
 KW attention disorder; cognition disorder; hypotension; hypertension;
 KW psychotic disorder; neurological disorder; dyskinesia;
 KW metabolic disorder; organ transplant rejection; enzyme; gene; ds.
 OS Homo sapiens.
 XX
 XX WO2003057841-A2.
 PN 17-JUL-2003.
 XX
 XX 31-DEC-2002; 2002WO-US041687.
 XX
 XX 31-DEC-2001; 2001US-0343169P.
 PR (GRIG/) GRIGORIEV I V.
 PA (SUDA/) SUDARSANAM S.
 XX
 XX Grigoriev IV, Sudarsanam S;
 PI
 XX WPI; 2003-587115/55.
 DR
 XX
 XX New isolated, enriched or purified nucleic acid molecule encoding a
 PT kinase polypeptide, useful for treating cancer, immune-related diseases,
 PT cardiovascular disease, brain or neuronal-associated diseases and
 PT metabolic disorders.
 XX
 PS Claim 33; SEQ ID NO 102; 491pp; English.

XX The invention relates to novel isolated, enriched or purified nucleic acid
 CC molecules encoding a kinase polypeptide. The nucleic acid molecule
 CC comprises a sequence that: (a) encodes a kinase polypeptide; (b) is a
 CC complement of (a); (c) hybridizes under stringent conditions to (a) and
 CC encodes a naturally occurring kinase polypeptide; (d) encodes the
 CC polypeptide in (a), except that it lacks one or more, but not all, of an
 CC N-terminal domain, C-terminal catalytic domain, a catalytic region and a C-
 CC terminal domain, a coiled-coil structure region, a spacer region and a C-
 CC terminal tail; or (e) is a complement of (d). The nucleic acid molecules,
 CC polypeptides, methods and substance are useful for treating cancers, brain or
 CC immune-related diseases or disorders, cardiovascular disease, brain or
 CC neuronal-associated diseases, and metabolic disorders. The disorders are
 CC preferably cancers of the tissues or of hematopoietic origin, diseases of
 CC the central or peripheral nervous system, Alzheimer's disease,
 CC Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis,
 CC viral infections, infections caused by prions, infections caused by
 CC bacteria, infections caused by fungi, ocular diseases, migraines, pain,
 CC sexual dysfunction, mood disorders, attention disorders, cognition
 CC disorders, hypotension, hypertension, psychotic disorders, neurological
 CC disorders, dyskinesias, metabolic disorders and organ transplant
 CC rejection. This sequence corresponds to the DNA encoding one of the
 CC kinase polypeptides of the invention.

XX SQ Sequence 2099 BP; 484 A; 592 C; 628 G; 395 T; 0 U; 0 Other;

Query Match 98.8%; Score 1447; DB 10; Length 2099;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1458; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

4 GGTGGTGTGGTGAAGAGATGGAGTCTCTGCTGCTCCACCTCCAGCTCGCGTAGG 63
 101 GGTGGTGTGGTGAAGAGATGGAGTCTCTGCTGCTCCACCTCCAGCTCGCGTAGG 160
 64 GTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 123
 161 GTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 220
 124 GTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 183
 221 GTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 280
 184 CTCAATATACAGTTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 243
 281 CTCAATATACAGTTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 340
 244 GGTTCAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 303
 341 GGTTCAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 400
 304 TATGGACCCACAGAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 363
 401 TATGGACCCACAGAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 460
 364 GTGATCCCGGGGCTCTCATGGACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 423
 461 GTGATCCCGGGGCTCTCATGGACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 520
 424 CGGCAATGGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483
 521 CGGCAATGGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 580
 484 GACCTCTGGACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
 581 GACCTCTGGACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
 544 CTGATTCGGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 603
 641 CTGATTCGGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700
 604 CTGCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
 701 CTGCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 760

Qy 664 AGTCATGCTGTGCTCTGCTGCTCAAGTGTGACGAGCGGGAAGCTTTGGCCCATTTCCGCG 723
 Db 761 AGTCATGCTGTGCTCTGCTGCTCAAGTGTGACGAGCGGGAAGCTTTGGCCCATTTCCGCG 820
 Qy 724 CGAGAGGAAACCTCTACCTGATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 783
 Db 821 CGAGAGGAAACCTCTACCTGATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 880
 Qy 784 AACAGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 843
 Db 881 AACAGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 940
 Qy 844 AAAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 903
 Db 941 AAAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1000
 Qy 904 ACTGCGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 963
 Db 1001 ACTGCGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1060
 Qy 964 GCGCTGTGAGAGCGCTTCTACCTAGACACAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1023
 Db 1061 GCGCTGTGAGAGCGCTTCTACCTAGACACAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1120
 Qy 1024 AAGGAGGTGATCAATCGGCTTCTTACCAATGAGAGCTGCTGCTGCTGCTGCTGCTGCT 1083
 Db 1121 AAGGAGGTGATCAATCGGCTTCTTACCAATGAGAGCTGCTGCTGCTGCTGCTGCTGCT 1180
 Qy 1084 GTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGAGCGCGAGGCCCT 1143
 Db 1181 GTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGAGCGCGAGGCCCT 1240
 Qy 1144 GAGGAGAGGAGATTTGGAGCGCTTGGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAGCCCT 1203
 Db 1241 GAGGAGAGGAGATTTGGAGCGCTTGGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAGCCCT 1300
 Qy 1204 CTCAGCCCTTACAGAACTAAGCAGCATGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAG 1263
 Db 1301 CTCAGCCCTTACAGAACTAAGCAGCATGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAG 1360
 Qy 1264 TTGGAACCTGTGCTTGGCTTCCAGGGAGCGAGGGGGGGGGGGGGGGGGGGGGGGGG 1323
 Db 1361 TTGGAACCTGTGCTTGGCTTCCAGGGAGCGAGGGGGGGGGGGGGGGGGGGGGGGGG 1420
 Qy 1324 CGAGAGCGGATGCTGCTAATGAAGACAGTAGAAGAGAGAGGAGGCTAG-AGATTGAGAG 1382
 Db 1421 CGAGAGCGGATGCTGCTAATGAAGACAGTAGAAGAGAGAGGAGGCTAGTAGATTGAGAG 1480
 Qy 1383 TAAGACGAGCAAGCAAAAGAACTGAGGGCCCAAGATGTTGGCCCAAGAGGCTGAGGAAAG 1442
 Db 1481 TAAGACGAGCAAGCAAAAGAACTGAGGGCCCAAGATGTTGGCCCAAGAGGCTGAGGAAAG 1540
 Qy 1443 GAACCAATTTGCCCAATG 1461
 Db 1541 GAACCAATTTGCCCAATG 1559

RESULT 7

AAF15853
 ID AAF15853 standard; cDNA; 2104 BP.
 XX
 AC AAF15853;
 DT
 XX 13-MAR-2001 (first entry)
 XX Human prostate cancer antigen nucleotide sequence SEQ ID NO:288.
 DE Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
 KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
 KW vulnary; gastrointestinal; nephrotropic; anti-infective; gynaecological;
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;

KW wound; infectious disease; ss.
 OS Homo sapiens.
 XX WO200055174-A1.
 XX 21-SEP-2000.
 XX 08-MAR-2000; 200WO-US005988.
 XX 12-MAR-1999; 99US-0124270P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX (ROSE/) ROSEN C A.
 XX Rosen CA, Ruben SM;
 XX WPI: 2000-587513/55.
 XX P-PSDB; AAB56650.
 XX
 XX Prostate cancer associated gene sequences, referred to as prostate cancer
 XX antigens, useful for treatment, prevention, and diagnosis of disorders
 XX such as prostate cancer.
 XX
 XX Claim 1; Page 805-806; 2338pp; English.
 XX
 XX AAF15566 to AAF16505 encode the human prostate cancer associated
 XX proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
 XX The prostate cancer antigens can have neuroprotective, cytostatic,
 XX cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
 XX nephrotropic, anti-infective, gynecological and antibacterial activities,
 XX and can be used in gene therapy. The prostate cancer antigen
 XX polynucleotides may be used for detection of prostate cancer, chromosome
 XX identification, as chromosome markers, and for numerous other diagnostic
 XX or research purposes. The prostate cancer antigens may be used to treat
 XX disorders such as neutral, immune, muscular, reproductive,
 XX gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 XX disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
 XX AAB57303 represent sequences used in the exemplification of the present
 XX invention
 XX
 XX Sequence 2104 BP; 517 A; 580 C; 597 G; 400 T; 0 U; 10 Other;
 XX
 XX Query Match 98.6%; Score 1443.2; DB 3; Length 2104;
 XX Best Local Similarity 99.4%; Pred. No. 0;
 XX Matches 1450; Conservative 7; Mismatches 1; Indels 1; Gaps 1;
 XX
 XX 4 GGTGGCTGTGGCTAAGCAAGATTGGAGCTACTGCTGTCACCTCCAGCTCGCGTAAGG 63
 XX 67 GGTGGCTGTGGCTAAGCAAGATTGGAGCTACTGCTGTCACCTCCAGCTCGCGTAAGG 126
 XX
 XX 64 GTGGCTGTGGCTAAGCAAGATTGGAGCTACTGCTGTCACCTCCAGCTCGCGTAAGG 123
 XX 127 GTGGCTGTGGCTAAGCAAGATTGGAGCTACTGCTGTCACCTCCAGCTCGCGTAAGG 186
 XX
 XX 124 GTGGGGGCGATGGACAGCTGCTCTAGAGATTGCTAACTGGAGAACACACAGGAGACT 183
 XX 187 GTGGGGGCGATGGACAGCTGCTCTAGAGATTGCTAACTGGAGAACACACAGGAGACT 246
 XX
 XX 184 CTCMAATACAGTTTGTATGCTTCTATGGGAGAGGAGTACTAGCAGGACATCTATGCA 243
 XX 247 CTCMAATACAGTTTGTATGCTTCTATGGGAGAGGAGTACTAGCAGGACATCTATGCA 306
 XX
 XX 244 GGTTCAGTCAGCCCATCTTAAGCACTTCTGGAAGGGCAGATGCCAGTGTGCTTGGC 303
 XX 307 GGTTCAGTCAGCCCATCTTAAGCACTTCTGGAAGGGCAGATGCCAGTGTGCTTGGC 366
 XX
 XX 304 TATGGACCCACAGAGCTGGGAAAGACGACACATGCTGGGCGACCCAGAGCAACCTGGG 363
 XX 367 TATGGACCCACAGAGCTGGGAAAGACGACACATGCTGGGCGACCCAGAGCAACCTGGG 426
 XX
 XX 364 GTGATCCCGGGCTCTCATGGACCTCTCTGAGCTCAAGAGGAGGAGGTGCGAGGGC 423

Db 427 GTGATCCGCGGGCTCTCATGAGACCTCTCTGAGCTCAAGAGGAGGAGGCTGCCAGGGC 486
 Qy 424 CGGCATGGGCGCTTTCTGTCAACCATGCTTTACCTAGAGATCTACACAGAGAGGTATTA 483
 Db 487 CGGCATGGGCGCTTTCTGTCAACCATGCTTTACCTAGAGATCTACACAGAGAGGTATTA 546
 Qy 484 GACCTCTGGACCTCTGCTTCCGAGAGACTGCTGTAATCCGAGAGAGTGCCTGGGGAATATC 543
 Db 547 GACCTCTGGACCTCTGCTTCCGAGAGACTGCTGTAATCCGAGAGAGTGCCTGGGGAATATC 606
 Qy 544 CTGATTCGGGCTCTCTCCAGAGCCCATGAGTACTGCTGTAATCCGAGAGAGTGCCTGGGGAATATC 603
 Db 607 CTGATTCGGGCTCTCTCCAGAGCCCATGAGTACTGCTGTAATCCGAGAGAGTGCCTGGGGAATATC 666
 Qy 604 CTGCAGCCAGTCCGAAATCGAGTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGCG 663
 Db 667 CTGCAGCCAGTCCGAAATCGAGTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGCG 726
 Qy 664 AGTCATGCTGTGCTCTGTCCTGTCAGGTCGAGCGGGAACGTTTGGCCCTCATTTGCGCAG 723
 Db 727 AGTCATGCTGTGCTCTGTCCTGTCAGGTCGAGCGGGAACGTTTGGCCCTCATTTGCGCAG 786
 Qy 724 CGAGAGGAAAACCTCTACCTGATTTGACTTGGCTGAGGAGCAACCGGCGCACAGGC 783
 Db 787 CGAGAGGAAAACCTCTACCTGATTTGACTTGGCTGAGGAGCAACCGGCGCACAGGC 846
 Qy 784 AACAGGGCTTCCGGCTAAAAGAGTGGAGCCATCAACCTCCCTGTTTGTCTGGGC 843
 Db 847 AACAGGGCTTCCGGCTAAAAGAGTGGAGCCATCAACCTCCCTGTTTGTCTGGGC 906
 Qy 844 AAAGTGTAGATGCGCTGAATCAGGGCTCTCTGCTGTACCTTATCGGAGCAGCAAGCTC 903
 Db 907 AAAGTGTAGATGCGCTGAATCAGGGCTCTCTGCTGTACCTTATCGGAGCAGCAAGCTC 966
 Qy 904 ACTGCGCTATTCAGGACTCTCTGGTGGCTCAGCCACAGTATCTTATTCGCAACAT 963
 Db 967 ACTGCGCTATTCAGGACTCTCTGGTGGCTCAGCCACAGTATCTTATTCGCAACAT 1026
 Qy 964 GCCCTCAGAGAGCTTCTACCTAGACACAGTCTCCGCTCAGCTCAACTTGTGTCAGGTC 1023
 Db 1027 GCCCTCAGAGAGCTTCTACCTAGACACAGTCTCCGCTCAGCTCAACTTGTGTCAGGTC 1086
 Qy 1024 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTGAGCCCTCATGCTTGGAGCT 1083
 Db 1087 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTGAGCCCTCATGCTTGGAGCT 1146
 Qy 1084 GTTAAGCTGTCTCAGAAAGAAATTGCTTGGTCCACAGAGGCAAGAGAGCCGAGGCT 1143
 Db 1147 GTTAAGCTGTCTCAGAAAGAAATTGCTTGGTCCACAGAGGCAAGAGAGCCGAGGCT 1206
 Qy 1144 GAGGAGAGGAGATGGGAGCCCTGAGCCCATGAGCCCTGAGCCCTGAGCCCTGAGCCCT 1203
 Db 1207 GAGGAGAGGAGATGGGAGCCCTGAGCCCATGAGCCCTGAGCCCTGAGCCCTGAGCCCT 1266
 Qy 1204 CTCAGCCCTTACAGAGCTAAGCAGCATGAGCCCGGCTTGTGGAGCGCTCTCT -CAG 1262
 Db 1267 CTCAGCCCTTACAGAGCTAAGCAGCATGAGCCCGGCTTGTGGAGCGCTCTCT -CAG 1326
 Qy 1263 CTTGGACCGCTGCTTGTGCTTCCAGGGGAGCCAGGGGGCCCTCTGTTGAGTACCCAAA 1322
 Db 1327 CTTGGACCGCTGCTTGTGCTTCCAGGGGAGCCAGGGGGCCCTCTGTTGAGTACCCAAA 1386
 Qy 1323 GCGAGAGCGGATGGTCTAATGAAGCAGTAGAGAGAGAGCCTAGAGATTGAGGCT 1382
 Db 1387 GCGAGAGCGGATGGTCTAATGAAGCAGTAGAGAGAGAGCCTAGAGATTGAGGCT 1446
 Qy 1383 TAAGAGCAAGAAAAAGAACTGAGGCGCAAGATGTTGGCCAGAGGCTGAGGAAAAAGGA 1442
 Db 1447 TAAGAGCAAGAAAAAGAACTGAGGCGCAAGATGTTGGCCAGAGGCTGAGGAAAAAGGA 1506
 Qy 1443 GAACCAATTGTCCCAATG 1461
 Db 1507 GAACCAATTGTCCCAATG 1525

XX DT 23-SEP-2004 (first entry)
 XX DE Human microtubule motor protein DNA #3.
 XX KW Human; microtubule motor protein; gene; ds;
 XX KW cellular proliferation disorder; cancer; hyperplasia; restenosis;
 KW cardiac hypertrophy; immune disorder; inflammation;
 KW kinesin-like DNA binding protein; KID; autoimmune disease; arthritis;
 KW graft rejection; inflammatory bowel disease; angioplasty.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FT 1. .1041
 XX FT CDS /*tag= a
 FT /product= "Human microtubule motor protein #3"
 XX US6762043-B1.
 XX 13-JUL-2004.
 XX 06-MAR-2002; 2002US-00093317.
 XX 20-APR-1999; 99US-00295612.
 XX 20-JUN-2000; 2000US-00597292.
 XX 28-NOV-2000; 2000US-00724224.
 XX (CYTO-) CYTOKINETICS INC.
 XX Beraud C;
 XX WPI; 2004-532491/51.
 XX P-PSDB; ADQ60232.
 XX New isolated microtubule motor protein, useful for screening modulators
 PT for treating cellular proliferation disorders such as cancer,
 PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
 PT inflammation.
 XX Disclosure; SEQ ID NO 5; 26pp; English.
 XX The invention relates to human microtubule motor proteins and the nucleic
 CC acids encoding them. The invention also relates to a method of screening
 CC for modulators of a motor protein which has microtubule stimulated ATPase
 CC activity, a method of testing for ATPase activity of microtubule motor
 CC proteins, methods to identify candidate agents that bind to a target
 CC protein or act as a modulator of the binding characteristics or
 CC biological activity of a target protein, modulators of the target
 CC protein, and methods of treating cellular proliferation disorders such as
 CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
 CC and inflammation, for treating disorders associated with kinesin-like DNA
 CC binding protein (KID) and for inhibiting KID. The sequences are used for
 CC screening for modulators of motor proteins useful for treating cellular
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
 CC hypertrophy, immune disorders and inflammation, for treating disorders
 CC associated with KID and for inhibiting KID and for treating autoimmune
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and
 CC proliferation induced after medical procedures including surgery and
 CC angioplasty. This sequence represents DNA encoding a human microtubule
 CC motor protein of the invention.
 XX SQ Sequence 1041 BP; 228 A; 295 C; 302 G; 216 T; 0 U; 0 Other;
 Query Match 70.9%; Score 1038.4; DB 12; Length 1041;
 Best Local Similarity 99.9%; Pred. No. 3.1e-278;
 Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGGGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCAGCTCGGTA 60
 Db 1 ATGGGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCAGCTCGGTA 60
 QY 61 AGGGTGGCTGTGGCACTGCGGCCATTGTTGGATGGAAACAGCGGAGCAAGTGATCCCCC 120

Db 61 AGGGTGGCTGTGGCACTGCGGCCATTGTTGGATGGAAACAGCGGAGCAAGTGATCCCCC 120
 QY 121 TGTGTGGGGGCGATGGACAGCTGCTCTCTAGAGATTGCTAATCGAGAAACACACAGAG 180
 Db 121 TGTGTGGGGGCGATGGACAGCTGCTCTCTAGAGATTGCTAATCGAGAAACACACAGAG 180
 QY 181 ACTCTCAATACCACTGTTGATGCTTCTATGGGAGAGAGTACTCTACGACGACATCTAT 240
 Db 181 ACTCTCAATACCACTGTTGATGCTTCTATGGGAGAGAGTACTCTACGACGACATCTAT 240
 QY 241 GCAGGTTCACTGTCAGCCCATCTTAAGGCACCTTCTGGAAGGCGAGAATGCCAGTGTGCT 300
 Db 241 GCAGGTTCACTGTCAGCCCATCTTAAGGCACCTTCTGGAAGGCGAGAATGCCAGTGTGCT 300
 QY 301 GCCTATGGACCCACAGGAGCTGGGAAGACGACACCAATGCTGGGCGCCACAGACAACCT 360
 Db 301 GCCTATGGACCCACAGGAGCTGGGAAGACGACACCAATGCTGGGCGCCACAGACAACCT 360
 QY 361 GGGGTGATCCCGGGGCTCTCATGGACCTCTCTGAGCTCAACAGGAGGAGGGTGCAG 420
 Db 361 GGGGTGATCCCGGGGCTCTCATGGACCTCTCTGAGCTCAACAGGAGGAGGGTGCAG 420
 QY 421 GGGCGGCATGGGCCCTTTCTGTCAACATGTTTACCTAGAGATCTTACAGAGAAGGTA 480
 Db 421 GGGCGGCATGGGCCCTTTCTGTCAACATGTTTACCTAGAGATCTTACAGAGAAGGTA 480
 QY 481 TTAGACCTCTGGACCCCTGCTTGGGAGACCTGTTAATCCGAGAAGACTGCGGGGGGAT 540
 Db 481 TTAGACCTCTGGACCCCTGCTTGGGAGACCTGTTAATCCGAGAAGACTGCGGGGGGAT 540
 QY 541 ATCTGATTCGGGTCTCTCCAGAACCCATCAGTAGCTTGTGCTGATTTTGGCGGCAC 600
 Db 541 ATCTGATTCGGGTCTCTCCAGAACCCATCAGTAGCTTGTGCTGATTTTGGCGGCAC 600
 QY 601 TTCTGTCACCCAGTGGAAATCGGACTGTAGGAGCCACCCGGCTCAACACGCGCTCTCC 660
 Db 601 TTCTGTCACCCAGTGGAAATCGGACTGTAGGAGCCACCCGGCTCAACACGCGCTCTCC 660
 QY 661 CGCAGTCATGCTGTGCTCTGCTCAAGGTGGACCGCGGAGCGTTTGGCCCCATTTCCG 720
 Db 661 CGCAGTCATGCTGTGCTCTGCTCAAGGTGGACCGCGGAGCGTTTGGCCCCATTTCCG 720
 QY 721 CAGCAGAGGAGGAAACTCTACTGATTGACTTGGCTGGGTCAGAGGACAAACCGCGCACA 780
 Db 721 CAGCAGAGGAGGAAACTCTACTGATTGACTTGGCTGGGTCAGAGGACAAACCGCGCACA 780
 QY 781 GGCAACAAGGGCCTTGGCTTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTGCTG 840
 Db 781 GGCAACAAGGGCCTTGGCTTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTGCTG 840
 QY 841 GGCAAGTGTAGATGGGCTGAATCAGGGCTCTCTGCTGTACTCTTATCGGACAGCAAG 900
 Db 841 GGCAAGTGTAGATGGGCTGAATCAGGGCTCTCTGCTGTACTCTTATCGGACAGCAAG 900
 QY 901 CTCACCTGCTCTATTGAGGAGCTCTGCTGGTGGCTCAGCCCAACATCTTTATTGCCAAC 960
 Db 901 CTCACCTGCTCTATTGAGGAGCTCTGCTGGTGGCTCAGCCCAACATCTTTATTGCCAAC 960
 QY 961 ATTGCCCCCTGAGAGAGCTTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGG 1020
 Db 961 ATTGCCCCCTGAGAGAGCTTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGG 1020
 QY 1021 TCCAAGGAGTGATCAATCG 1040
 Db 1021 TCCAAGGAGTGATCAATCG 1040

RESULT 10
 ADC23337
 ID ADC23337 standard; DNA; 1115 BP.
 XX
 AC ADC23337;

Fri Nov 12 12:26:24 2004

XX	18-DEC-2003	(first entry)	Db	198	GTGCGGGCATGGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACCAACGAGGAGACT	257
DT						
XX		DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 1).	Qy	184	CTCAATACACAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	243
DE						
XX		human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;	Db	258	CTCAATACACAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	317
KW		cytostatic; cardiant; immunomodulator; antiinflammatory; gene therapy;	Qy	244	GGTTCAGTGCAGGCCATCTCTAAGGCACCTTCTGAGGGGAGGAGTACTCAGTGTCTTGGCC	303
KW		cancer; hyperplasia; restenosis; cellular proliferation disorder;	Db	318	GGTTCAGTGCAGGCCATCTCTAAGGCACCTTCTGAGGGGAGGAGTACTCAGTGTCTTGGCC	377
XX		cardiac hypertrophy; immune disorder; inflammation.				
OS		Homo sapiens.	Qy	304	TATGAGCCCAACAGGAGCTGGGAGAGCACAATATGCTGGGAGCCAGAGCAACCTGGG	363
XX						
XX		Location/Qualifiers	Db	378	TATGAGCCCAACAGGAGCTGGGAGAGCACAATATGCTGGGAGCCAGAGCAACCTGGG	437
FH		1. .1115				
FT		/*tag= a	Qy	364	GTGATCCCGCGGGCTCTCATGGACCTCTGAGCAGCTCACAGGAGGAGGGTGGCCGAGGGC	423
FT		/product= "KID protein"	Db	438	GTGATCCCGCGGGCTCTCATGGACCTCTGAGCAGCTCACAGGAGGAGGGTGGCCGAGGGC	497
FT		/transl_except= (pos: 1. .5; aa: Met)	Qy	424	CGGCAATGGGCCCCCTTCTGTCTACCATGTCTTACCTAGAGATCTACAGGAGGAGGATTA	483
FT		/note= "this codon has an apparent 2 nucleotide insertion	Db	498	CGGCAATGGGCCCCCTTCTGTCTACCATGTCTTACCTAGAGATCTACAGGAGGAGGATTA	557
FT		that alters the reading frame"				
XX		US6387644-B1.	Qy	484	GACCTCTCTGGAACCTGTCTGGGAGACCTGTGTAATCCGAGAGACTGCGCGGGGAATATC	543
XX						
XX		14-MAY-2002.	Db	558	GACCTCTCTGGAACCTGTCTGGGAGACCTGTGTAATCCGAGAGACTGCGCGGGGAATATC	617
XX		28-NOV-2000; 2000US-00724224.	Qy	544	CTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTTGGCGGCACATTC	603
XX		20-APR-1999; 99US-00295612.	Db	618	CTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTTGGCGGCACATTC	677
PR		20-JUN-2000; 2000US-00597292.	Qy	604	CTGCCAGGCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCCCGC	663
XX		(CYTO-) CYTOKINETICS INC.	Db	678	CTGCCAGGCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCCCGC	737
XX		Beraud C;	Qy	664	AGTCATGTGTGCTCTCTGTCAGGTGCAAGGTGCAAGCGGGGAAACGTTTGGCCCCCATTTGCCAG	723
XX		WPI; 2003-706919/67.	Db	738	AGTCATGTGTGCTCTCTGTCAGGTGCAAGGTGCAAGCGGGGAAACGTTTGGCCCCCATTTGCCAG	797
DR		P-PSDB; ADC23338.	Qy	724	CGAGAGGGGAAACCTCTACCTGATGACTTGGCTGGGTGAGAGGACAAACCGCGCAGCAGGC	783
XX		Identifying a candidate agent as modulator of function of a target	Db	798	CGAGAGGGGAAACCTCTACCTGATGACTTGGCTGGGTGAGAGGACAAACCGCGCAGCAGGC	857
PT		protein for treating cellular proliferation disorders by adding a	Qy	784	AAACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCTGTTTGTCTCTGGGC	843
PT		candidate agent to a mixture of the target protein that	Db	858	AAACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCTGTTTGTCTCTGGGC	917
PT		directly/indirectly produces ADP or phosphate.	Qy	844	AAAGTGTAGATGCGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGGACAGCAAGCTC	903
XX		Disclosure; SEQ ID NO 1; 26pp; English.	Db	918	AAAGTGTAGATGCGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGGACAGCAAGCTC	977
XX			Qy	904	ACTCGCTATTTCAGGACTCTCTGGGTGGCTCAGCCCAACAGTATCTTATTCGCAACATTT	963
CC		This invention relates to a novel method for high throughput screening	Db	978	ACTCGCTATTTCAGGACTCTCTGGGTGGCTCAGCCCAACAGTATCTTATTCGCAACATTT	1037
CC		systems used to identify compounds for the treatment of cellular	Qy	964	GCCTCTGAGAGAGCGTTTCTACCTAGACAGCTCTCCGCACTCAACTTCTGCTGCAGGCTC	1023
CC		proliferation disorders. Specifically, it refers to candidate agents that	Db	1038	GCCTCTGAGAGAGCGTTTCTACCTAGACAGCTCTCCGCACTCAACTTCTGCTGCAGGCTC	1097
CC		are capable of modulating the activity of target proteins having motor	Qy	1024	AGGAGGTGATCAATCG	1040
CC		domains, such that the target protein directly or indirectly produces ADP	Db	1098	AAGGAGGTGATCAATCG	1114
CC		or phosphate. Furthermore, this activity can be determined using				
CC		fluorescence or absorbance readouts. The present invention describes a				
CC		method that identifies modulators of the target protein, which is a				
CC		kinesin-like DNA binding protein (known as KID) as cytostatic, cardiants,				
CC		immunomodulators and antiinflammatories. Accordingly, through gene				
CC		therapy, they can be used for the treatment of cancer, hyperplasias,				
CC		restenosis, cardiac hypertrophy, immune disorders and inflammation. This				
CC		polynucleotide sequence is human KID DNA (SeqID 1) encoding a full length				
CC		KID enzyme of the invention.				
XX			Qy	Sequence 1115 BP; 241 A; 318 C; 333 G; 223 T; 0 U; 0 Other;		
XX						
XX		Query Match		70.7%; Score 1035.4; DB 10; Length 1115;		
XX		Best Local Similarity		99.9%; Pred. No. 2.2e-277;		
XX		Matches 1036; Conservative		0; Mismatches 1; Indels 0; Gaps 0;		
Qy		4 GGTGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGTGTCACCTCCAGCTCGGTAAGG	63			
Db		78 GGTGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGTGTCACCTCCAGCTCGGTAAGG	137			
Qy		64 GTGGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGTGTCACCTCCAGCTCGGTAAGG	123			
Db		138 GTGGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGTGTCACCTCCAGCTCGGTAAGG	197			
Qy		124 GTGGGGGCGATGGAGCTGCTCTCTAGAGATTGCTAAGTGGAGGAACCAACGAGGAGCT	183			

RESULT 11
ADQ60227 standard; DNA; 1115 BP.
ID ADQ60227
XX
AC ADQ60227;
XX
DT 23-SEP-2004 (first entry)
XX
DE Human microtubule motor protein DNA #1.

XX Human; microtubule motor protein; gene; ds;
 KW cellular proliferation disorder; cancer; hyperplasia; restenosis;
 KW cardiac hypertrophy; immune disorder; inflammation;
 KW kinesin-like DNA binding protein; KID; autoimmune disease; arthritis;
 KW graft rejection; inflammatory bowel disease; angioplasty.
 XX
 OS Homo sapiens.
 XX
 FN US6762043-B1.
 XX
 PD 13-JUL-2004.
 XX
 PF 06-MAR-2002; 2002US-00093317.
 XX
 PR 20-APR-1999; 99US-00295612.
 PR 20-JUN-2000; 2000US-00597292.
 PR 28-NOV-2000; 2000US-00724224.
 XX
 PA (CYTO-) CYTOKINETICS INC.
 XX
 PI Beraud C;
 XX
 WI 2004-532491/51.
 XX
 DR New isolated microtubule motor protein, useful for screening modulators
 PT for treating cellular proliferation disorders such as cancer,
 PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
 PT inflammation.
 XX
 PS Disclosure; SEQ ID NO 1; 26pp; English.
 XX
 CC The invention relates to human microtubule motor proteins and the nucleic
 CC acids encoding them. The invention also relates to a method of screening
 CC for modulators of a motor protein which has microtubule stimulated ATPase
 CC activity, a method of testing for ATPase activity of microtubule motor
 CC proteins, methods to identify candidate agents that bind to a target
 CC protein or act as a modulator of the binding characteristics or
 CC biological activity of a target protein, modulators of the target
 CC protein, and methods of treating cellular proliferation disorders such as
 CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
 CC and inflammation, for treating disorders associated with kinesin-like DNA
 CC binding protein (KID) and for inhibiting KID. The sequences are used for
 CC screening for modulators of motor proteins useful for treating cellular
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
 CC hypertrophy, immune disorders and inflammation, for treating disorders
 CC associated with KID and for inhibiting KID and for treating autoimmune
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and
 CC proliferation induced after medical procedures including surgery and
 CC angioplasty. This sequence represents DNA encoding a human microtubule
 CC motor protein of the invention. Note: The specification states that this
 CC sequence encodes the protein featured as SEQ ID NO:2, but this does not
 CC appear to be the case.
 XX
 SQ Sequence 1115 BP; 241 A; 318 C; 333 G; 223 T; 0 U; 0 Other;
 Query Match 70.7%; Score 1035.4; DB 12; Length 1115;
 Best Local Similarity 99.9%; Pred. No. 2.2e-277;
 Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 GTGCGCTGTGCGGTAAGCAAGATTGGAGTACTCGTCGTCCACCTCCAGCTCGCGTAAGG 63
 DB 78 GGTGCTGTGCGGTAAGCAAGATTGGAGTACTCGTCGTCCACCTCCAGCTCGCGTAAGG 137
 QY 64 GTGCGCTGTGCGGTAAGCAAGATTGGAGTACTCGTCGTCCACCTCCAGCTCGCGTGT 123
 DB 138 GTGCGCTGTGCGGTAAGCAAGATTGGAGTACTCGTCGTCCACCTCCAGCTCGCGTGT 197
 QY 124 GTGCGGCGCATGGACAGCTGCTCTCTAGAGATTGCTTAACCTGGAGAACCAACAGGAGACT 183
 DB 198 GTGCGGCGCATGGACAGCTGCTCTCTAGAGATTGCTTAACCTGGAGAACCAACAGGAGACT 257
 QY 184 CTCAAATACCAGTTTGTATGCTTCTATGCGGAGAGGAGTACTCAGCAGGAGACATCTATGCA 243

Db 258 CTCAAATACCAGTTTGTATGCTTCTATGCGGAGAGGAGTACTCAGCAGGACATCTATGCA 317
 QY 244 GGTTCAGTGCAGGCCCATCTTAAGGCACCTTGTGGAAGGCGAGAAATGCCAGTGTGCTTGC 303
 DB 318 GGTTCAGTGCAGGCCCATCTTAAGGCACCTTGTGGAAGGCGAGAAATGCCAGTGTGCTTGC 377
 QY 304 TATGGACCCACAGGAGCTGGGAAGACGACACAATGTCTGGGCAGCCACAGAGCAACCTGGG 363
 DB 378 TATGGACCCACAGGAGCTGGGAAGACGACACAATGTCTGGGCAGCCACAGAGCAACCTGGG 437
 QY 364 GTGATCCCGCGGGCTCTCATGGACCTCTCTGAGCTCAACAAGGAGAGGGTCCGAGGGC 423
 DB 438 GTGATCCCGCGGGCTCTCATGGACCTCTCTGAGCTCAACAAGGAGAGGGTCCGAGGGC 497
 QY 424 CGGCCATGGGCGCTTCTGTACCATCTCTTACCTAGAGATCTTACCAAGGAGAAAGTATTA 483
 DB 498 CGGCCATGGGCGCTTCTGTACCATCTCTTACCTAGAGATCTTACCAAGGAGAAAGTATTA 557
 QY 484 GACCTCTGGACCTCTCTGGGAGAGCTGTGTAATCCGAGAAAGTCTGCCGGGGAATATC 543
 DB 558 GACCTCTGGACCTCTCTGGGAGAGCTGTGTAATCCGAGAAAGTCTGCCGGGGAATATC 617
 QY 544 CTGATTCGGGCTCTCTCCAGAACCCCATCAGTAGCTTGTCTGATTTTGTAGCGGCATTC 603
 DB 618 CTGATTCGGGCTCTCTCCAGAACCCCATCAGTAGCTTGTCTGATTTTGTAGCGGCATTC 677
 QY 604 CTGCGACCCAGTCCGAATCCGAGCTGTAGGAGCCACCGGCTCAACCAAGGCGCTCTCCGCG 663
 DB 678 CTGCGACCCAGTCCGAATCCGAGCTGTAGGAGCCACCGGCTCAACCAAGGCGCTCTCCGCG 737
 QY 664 AGTCATCTCTGCTCTCTGGTCAAGGTGGACCGAGCGGAACGTTTGGCCCCCATTTCCGCCAG 723
 DB 738 AGTCATCTCTGCTCTCTGGTCAAGGTGGACCGAGCGGAACGTTTGGCCCCCATTTCCGCCAG 797
 QY 724 CGAGAGGAAAACCTTACTGATTGACTTGGCTGGGTGAGAGGACAAACCGCGGCACAGGC 783
 DB 798 CGAGAGGAAAACCTTACTGATTGACTTGGCTGGGTGAGAGGACAAACCGCGGCACAGGC 857
 QY 784 AACAGGGCCCTCGGCTTAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTCTGGGC 843
 DB 858 AACAGGGCCCTCGGCTTAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTCTGGGC 917
 QY 844 AAAGTGTAGATGCGCTGAATCAGGGGCTCTCTGCTGATCTTATCGGAGCAGCAAGCTC 903
 DB 918 AAAGTGTAGATGCGCTGAATCAGGGGCTCTCTGCTGATCTTATCGGAGCAGCAAGCTC 977
 QY 904 ACTGCGCTATTGAGAGACTCTCTGGGTGGCTCAGCCCAAGATATCTTATTCGCAACATT 963
 DB 978 ACTGCGCTATTGAGAGACTCTCTGGGTGGCTCAGCCCAAGATATCTTATTCGCAACATT 1037
 QY 964 GCCCCTGAGAGCGCTTCTTACCTAGACAGACTCTCCGACTCAACTTTGCTGCCAGGTCC 1023
 DB 1038 GCCCCTGAGAGCGCTTCTTACCTAGACAGACTCTCCGACTCAACTTTGCTGCCAGGTCC 1097
 QY 1024 AAGGAGGTGATCAATCG 1040
 DB 1098 AAGGAGGTGATCAATCG 1114

RESULT 12
 ABK70260
 ID ABK70260 standard; cDNA; 491 BP.
 XX
 AC ABK70260;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human lung cancer associated cDNA SEQ ID 131.
 XX
 KW Human; ss; gene; lung cancer; cytostatic; tumour; vaccine.
 XX
 OS Homo sapiens.

XX WO200224057-A2.
 XX 28-MAR-2002.
 XX 20-SEP-2001; 2001WO-US042232.
 XX 22-SEP-2000; 2000US-0234837P.
 XX 10-OCT-2000; 2000US-0239440P.
 XX 29-JUN-2001; 2001US-0301928P.
 XX (CORI-) CORIXA CORP.
 XX Benson DR, Mohamath R, Lodes MJ;
 XX WPI; 2002-372001/40:
 XX New tumor lung proteins and nucleic acids encoding the proteins, useful
 XX as vaccines and for treating, preventing, diagnosing or monitoring lung
 XX cancer.
 XX Claim 1; Page 144; 189pp; English.
 XX The invention relates to an isolated polynucleotide comprising a sequence
 XX selected from 183 human DNA sequences (appearing as ABK70130-ABK70312),
 XX or their fragments, homologues, variants or complements and their encoded
 XX polypeptides. Also included are an expression vector comprising the
 XX polynucleotide operably linked to an expression control sequence; a host
 XX cell transformed or transfected with an expression vector of; an isolated
 XX antibody, or its antigen-binding fragment that specifically binds to the
 XX polypeptide; a method for detecting the presence of a cancer in a patient
 XX ; a fusion protein comprising at least the polypeptide; an
 XX oligonucleotide that hybridises to the polynucleotide under moderately
 XX stringent conditions; a method for stimulating and/or expanding T cells
 XX specific for a tumor protein; an isolated T cell population comprising T
 XX cells prepared from the method of above; a composition comprising a first
 XX component consisting of carriers and immunostimulants, and a second
 XX component selected from the polynucleotides, proteins, antibodies, fusion
 XX proteins, T cell populations and antigen presenting cells expressing the
 XX polypeptide; methods for stimulating an immune response or treating
 XX cancer in a patient by administering the composition and diagnostic kits
 XX comprising at least one of the oligonucleotide of, or an antibody and a
 XX detection reagent consisting of a reporter group. The polypeptides and
 XX polynucleotides are useful as vaccines for the treatment or prevention of
 XX lung cancer, and for diagnosis and monitoring of such cancer. The
 XX polynucleotide, polypeptide and antigen presenting cells can be used to
 XX stimulate or expand T cells specific for a tumorous protein. The
 XX polynucleotides may be used as probes or primers for nucleic acid
 XX hybridisation, and in the preparation of ribozyme molecules for
 XX inhibiting expression of tumour polypeptides and proteins in tumour
 XX cells. The present sequence is one of the 183 lung cancer associated
 XX polynucleotides
 XX Sequence 491 BP; 127 A; 140 C; 136 G; 88 T; 0 U; 0 Other;
 Query Match 32.8%; Score 480; DB 6; Length 491;
 Best Local Similarity 99.8%; Pred. No. 5.3e-123;
 Matches 491; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 936 AGCCACAGATATCTTATGTCACATTCGCTGAGAGACGCTTCTACCTAGACACAGT 995
 DB 1 AGCCACAGATATCTTATGTCACATTCGCTGAGAGACGCTTCTACCTAGACACAGT 60
 QY 996 CTCGCGACTCACTTTGCTGCGAGTCCAGGAGTCAATCAATCGGCTTTTACCAGTGA 1055
 DB 61 CTCGCGACTCACTTTGCTGCGAGTCCAGGAGTCAATCAATCGGCTTTT-CCAAATGA 119
 QY 1056 GAGCTGCGAGCTCATGCTTGGACCTGTTAAGCTGCTCAGAAAGAAATGCTTGCTCC 1115
 DB 120 GAGCTGCGAGCTCATGCTTGGACCTGTTAAGCTGCTCAGAAAGAAATGCTTGCTCC 179
 QY 1116 ACCAGAGCAAGAGAGCCCGGCTTGGAGAAAGAGAGATTGGAGCCCTTGAGCCCAT 1175

DB 180 ACCAGAGCAAGAGAGCCCGGCTTGGAGAAAGAGAGATTGGAGCCCTTGAGCCCAT 239
 QY 1176 GGCAGCTCCAGCCCTCTGCTCCAGAAAACCTCAGCCCCCTACAGAGCTAAGCAGCATGGA 1235
 DB 240 GGCAGCTCCAGCCCTCTGCTCCAGAAAACCTCAGCCCCCTACAGAGCTAAGCAGCATGGA 299
 QY 1236 CCGGCCCATGCTGGAGCGCTCTCAGCTTGGAGCCGCTCTGCTTGCCTCCAGGGAGGCCA 1295
 DB 300 CCGGCCCATGCTGGAGCGCTCTCAGCTTGGAGCCGCTCTGCTTGCCTCCAGGGAGGCCA 359
 QY 1296 GGGGGCCCTCTGTTGAGTATACCCCAAGCGAGAGCGGATGGTGTCTAATGAAGCAGTAGA 1355
 DB 360 GGGGGCCCTCTGTTGAGTATACCCCAAGCGAGAGCGGATGGTGTCTAATGAAGCAGTAGA 419
 QY 1356 AGAGAAGACCTAGAGATTGAGAGCTTAAGACGAAGCAAAAGAACTGGAGGCCAAGAT 1415
 DB 420 AGAGAAGACCTAGAGATTGAGAGCTTAAGACGAAGCAAAAGAACTGGAGGCCAAGAT 479
 QY 1416 GTTGCCCCAGAA 1427
 DB 480 GTTGCCCCAGAA 491
 RESULT 13
 AAS38879
 ID AAS38879 standard; cDNA; 386 BP.
 XX AAS38879;
 XX 17-DEC-2001 (first entry)
 XX Novel human diagnostic and therapeutic gene #1937.
 XX Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.
 XX Homo sapiens.
 XX WO200166753-A2.
 XX 13-SEP-2001.
 XX 09-MAR-2001; 2001WO-US0007787.
 XX 09-MAR-2000; 2000US-0188609P.
 XX (CHIR) CHIRON CORP.
 XX (HYSE-) HYSEQ INC.
 XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
 XX Reinhard C, Randazzo F, Kennedy GC, Pot D, Kassam A, Lamson G;
 XX Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
 XX Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;
 XX WPI; 2001-530177/58.
 XX New polynucleotides and polypeptides, useful for diagnosis and treatment
 XX of breast, lung and colon cancer.
 XX Claim 1; Page 1080; 1193pp; English.
 XX The invention relates to new polynucleotides and polypeptides, useful for
 XX diagnosis and treatment of breast, lung and colon cancer. The sequences
 XX can be used in detecting differentially expressed genes correlated with a
 XX cancerous state of a mammalian cell, comprising detecting at least one
 XX differentially expressed gene product in a test sample derived from a
 XX cell suspected of being cancerous. They can also be used to inhibit
 XX tumour growth by modulating expression of a gene product. AAS36943-
 XX AAS39338 represent novel human diagnostic and therapeutic coding
 XX sequences of the invention
 XX Sequence 386 BP; 105 A; 102 C; 118 G; 61 T; 0 U; 0 Other;
 Query Match 25.7%; Score 376.4; DB 4; Length 386;

Best Local Similarity 98.4%; Pred. No. 3.2e-94;
Matches 380; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1073 CTTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTCCTTGGTCCACAGAGGCAAGAGAG 1132
Db |||||
1 CGTTGCTGCTGGTTAAGCTGTCTCAGAAAGAAATTCCTTGGTCCACAGAGGCAAGAGAG 60

QY 1133 CCCGAGGCCCTGAGGAAGAGAGATTGGGAGCCCTTGAGCCCATGCGAGCTCCAGCCTCTG 1192
Db |||||
61 CCCGAGGCCCTGAGGAAGAGAGATTGGGAGCCCTTGAGCCCATGCGAGCTCCAGCCTCTG 120

QY 1193 CCTCCCAAACTAGCCCTTACAGAGCTAAGCAGCATGAGCCCGCCATGCTGGAGC 1252
Db |||||
121 CCTCCCAAACTAGCCCTTACAGAGCTAAGCAGCATGAGCCCGCCATGCTGGAGC 180

QY 1253 GCCTCCTCAGCTTGACCGCTGCTGCTCCAGGGGAGCCAGGGGGCCCTCTGTTGA 1312
Db |||||
181 GCCTCCTCAGCTTGACCGCTGCTGCTCCAGGGGAGCCAGGGGGCCCTCTGTTGA 240

QY 1313 GTACCCCAAGCGAGCGGATGTTGTAATGAAGACAGTAGAAGAGAGGACCTAGAGA 1372
Db |||||
241 GTACCCCAAGCGAGCGGATGTTGTAATGAAGACAGTAGAAGAGAGGACCTAGAGA 300

QY 1373 TTGAGAGCTTAAAGACGAGCAAAAGAACTGGAGGCCCAAGATGTTGGCCCAAGAGCTG 1432
Db |||||
301 TTGAGAGCTTAAAGACGAGCAAAAGAACTGGAGGCCCAAGATGTTGGCCCAAGAGCTG 360

QY 1433 AGGAAAGGAGCAACCATGTCCTCCACA 1458
Db |||||
361 AGGAAAGGAGCAACCATGTCCTCCACA 386

RESULT 14
ACH44493
ID ACH44493 standard; cDNA; 464 BP.
XX ACH44493;
XX
XX
XX 13-OCT-2003 (first entry)
XX Human foetal brain cDNA #5218.
XX
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX genome mapping; biodiversity; genetic disorder.
XX Homo sapiens.
XX
XX US2003073623-A1.
XX
XX 17-APR-2003.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX (DRMA//) DRMANAC R T.
XX (LABA//) LABAT I.
XX (STAC//) STACHE-CRAIN B.
XX (DICK//) DICKSON M C.
XX (JONE//) JONES L W.
XX
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
XX mapping, in the recombinant production of protein, or in generating
XX antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 31705; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of

CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030073623
XX
SQ Sequence 464 BP; 96 A; 136 C; 131 G; 96 T; 0 U; 5 Other;

Query Match 18.1%; Score 265; DB 9; Length 464;
Best Local Similarity 77.6%; Pred. No. 3.5e-63;
Matches 333; Conservative 0; Mismatches 95; Indels 1; Gaps 1;

QY 126 GCGGGGCGATGGACAGCTGCTCTTAGAGATTGCTAAGTGGAGGAAACCAACGAGGACTCT 185
Db |||||
37 GCCTGGAATTCATATATCGGTCGATAGATGTAAACGGCCCGCAACCAACCGATAATCT 96

QY 186 CAATATACAGTTTATGCTTCTATGGGAGAGAGTACTCAGCAGAGCATCTATGCAGG 245
Db |||||
97 CCCTTACCTTGCCCTTGCCCTTTGACAGCCATAGAGGAGCTGCGCTTGACATGTACCTTCC 156

QY 246 TTCAGTGCAGCCCATCTTAAGGCACTTGTGGAAGGGCAGAAATGCCAGTGTGCTGCCTA 305
Db |||||
157 TTGACCGTGGACCGTGCCACTGCCACTGCTCTTCGTAAATAAGGAGGAGCTGGCCCTA 216

QY 306 TGGACCCACAGGAGCTGGGAAAGACGACACAATGCTGGGAGCCCGCAGAGCAACCTGGGGT 365
Db |||||
217 CGGTGCTGAGGAGCTGGCATTAC-ATCTGATGCTTTGGCAGCCCGCAGAGCAACCTGGGGT 275

QY 366 GATCCCGCGGGCTCTCATGGACCTCTGAGCTCACAAGGAGAGGGTGCAGAGGGCCG 425
Db |||||
276 GATCCCGCGGGCTCTCATGGACCTCTGAGCTCACAAGGAGAGGGTGCAGAGGGCCG 335

QY 426 GCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTTACCAGGAGAGGTATTAGA 485
Db |||||
336 GCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTTACCAGGAGAGGTATTAGA 395

QY 486 CCTCCTGGACCCCTGCTTGGGAGACCTGGTAATCCGAGAAAGACTGCGCGGGGAATATCCT 545
Db |||||
396 CCTCCTGGACCCCTGCTTGGGAGACCTTGGTAATCCGAGAAAGACTGCGCGGGGAATATCCT 455

QY 546 GATTCCGGG 554
Db |||||
456 GATTCCGGG 464

RESULT 15
ACH77408
ID ACH77408 standard; DNA; 531 BP.
XX ACH77408;
XX
XX 29-JUL-2004 (first entry)
XX
XX Human genome derived single exon probe #10603.
XX Human; probe; ss; gene expression; single exon probe; microarray;
XX alternative splicing event; genomic alteration.
XX
XX Homo sapiens.
XX
XX US2003194704-A1.
XX

Db 154 AGAGAGTGGAGCCATCAACACCTCCCTGTTGTTCTCGGGCAAGTGGTAGATGCGCTGAA 213
QY 864 TCAGGGCCCTCCCTCGTGTACCTTATCGGACAGCAAGCTCACTCGCCTATTGACAGAC 921
Db 214 TCAGGGCCCTCCCTCGTGTACCTTATCGGACAGCAAGCTCACTCGCCTATTGACAGTC 271

Search completed: November 10, 2004, 06:58:00
Job time : 760.977 secs

PD 16-OCT-2003.
XX 03-APR-2002; 2002US-00029386.
XX 03-APR-2002; 2002US-00029386.
XX (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANK/) HANKEL D K.
XX Penn SG, Rank DR, Hanzel DK;
XX WPI; 2004-119264/12.
XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX Claim 15; SEQ ID NO 10603; 80pp; English.
XX The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030194704
XX
XX SQ Sequence 531 BP; 117 A; 153 C; 149 G; 112 T; 0 U; 0 Other;
Query Match 15.9%; Score 233.2; DB 12; Length 531;
Best Local Similarity 98.7%; Pred. No. 2.6e-54;
Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 684 CAAGGTGGACACGCGGAACGTTTGGCCCATTTCCGACGAGAGGAAAACCTACCT 743
Db 34 CCAGGTGGACACGCGGAACGTTTGGCCCATTTCCGACGAGAGGAAAACCTACCT 93
QY 744 GATTGACTTGGCTGGGTACAGGACACCGCGCACAGGCAACAAGGCCCTTCGGCTAAA 803
Db 94 GATTGACTTGGCTGGGTACAGGACACCGCGCACAGGCAACAAGGCCCTTCGGCTAAA 153
QY 804 AGAGAGTGGAGCCATCAACACCTCCCTGTTGTTGCTCGGGCAAGTGGTAGATGCGCTGAA 863

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 ; Search time 75.9417 Seconds
(without alignments)
1747.786 Million cell updates/sec

Title: US-10-797-893-2

Perfect score: 1878

Sequence: 1 MAAGSTQQRREMAAASAA.....FYLDTVSALNFAARSKEVIN 370

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 359729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1878	100.0	370	7	Adc23338 Human kin
2	1878	100.0	370	8	Adg60228 Human mic
3	1878	100.0	490	7	Adk40973 Novel hum
4	1878	100.0	665	8	Adg09240 Human KNS
5	1873	99.7	512	7	Adc23340 Human kin
6	1873	99.7	512	8	Adg60230 Human mic
7	1763	93.9	346	8	Adc23342 Human kin
8	1763	93.9	346	8	Adg60232 Human mic
9	1763	93.9	460	3	Aab56650 Human pro
10	1763	93.9	487	7	Adc23344 Human kin
11	1763	93.9	487	8	Adg60234 Human mic
12	553.5	29.8	784	4	ABE71112 Drosophil
13	559	29.8	548	7	Adm04007 Human pro
14	559	29.8	998	8	Ado44167 Structura
15	552.5	29.4	357	5	Aau76958 Novel hum
16	552.5	29.4	898	4	ABU53123 Intracell
17	552.5	29.4	898	5	Aau76957 Novel hum
18	552.5	29.4	898	7	Adc31082 Human nov
19	552.5	29.4	898	7	ADI15915 Human pp
20	551.5	29.4	603	6	ABU11606 Human MDD
21	549	29.2	905	5	Aau76967 Novel hum
22	541	28.8	408	8	Adn40550 Candida a
23	541	28.8	408	8	Adn40554 Candida a
24	541	28.8	972	8	Adn40548 Candida a
25	541	28.8	974	8	Adn40552 Candida a

26	535.5	28.5	677	4	ABB65183 Drosophil
27	528	28.1	383	5	ABB07412 Amino aci
28	528	28.1	383	6	ABG72053 Human Hsk
29	528	28.1	864	5	ABB07410 Human kin
30	528	28.1	864	6	ABG72052 Human kin
31	526	28.0	375	5	AAE14401 Human Hsk
32	526	28.0	409	5	AAE14402 Human Hsk
33	526	28.0	409	5	Aau79592 Human Hsk
34	526	28.0	419	7	Adc31540 Human nov
35	526	28.0	1388	5	AAE14400 Human kin
36	526	28.0	1388	5	AAU79590 Human kin
37	526	28.0	1388	6	ABR48222 Human bla
38	526	28.0	1388	7	ADB80468 Ovarian c
39	526	28.0	1388	7	Adc35116 Human bre
40	526	28.0	1388	8	ADL83290 Human pro
41	526	28.0	1388	8	Adq20128 Human sof
42	526	28.0	1388	8	Adq09226 Human KNS
43	525.5	28.0	757	4	Aau19569 Human dia
44	525.5	28.0	757	5	ABP51294 Human MDD
45	523.5	27.9	834	7	ADJ95078 Novel NOV

ALIGNMENTS

RESULT 1
ADC23338
ID ADC23338 standard; protein; 370 AA.
XX
AC ADC23338;
XX
18-DEC-2003 (first entry)
DT
XX
DE Human kinesin-like DNA binding protein (KID) (SeqID 2).
XX

KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID;
KW cytosolic; cardiac; immunomodulator; antiinflammatory; gene therapy;
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;
KW cardiac hypertrophy; immune disorder; inflammation.
XX
OS Homo sapiens.
XX

FH Key Location/Qualifiers
FT Misc-difference 1
FT /note= "Encoded by ATGCA"
XX
PN US6387644-B1.
XX
PD 14-MAY-2002.
XX
PF 28-NOV-2000; 2000US-00724224.
XX
PR 20-APR-1999; 99US-00295612.
PR 20-JUN-2000; 2000US-00597292.
XX
PA (CYTO-) CYTOKINETICS INC.
XX
PI Beraud C;
XX
XX
DR WPI; 2003-706919/67.
DR N-PSDB; ADC23337.
XX

Identifying a candidate agent as modulator of function of a target protein for treating cellular proliferation disorders by adding a candidate agent to a mixture of the target protein that directly/indirectly produces ADP or phosphate.

Claim 1; SEQ ID NO 2; 26pp; English.

This invention relates to a novel method for high throughput screening systems used to identify compounds for the treatment of cellular proliferation disorders. Specifically, it refers to candidate agents that are capable of modulating the activity of target proteins having motor

CC domains, such that the target protein directly or indirectly produces ADP
CC or phosphate. Furthermore, this activity can be determined using
CC fluorescence or absorbance readouts. The present invention describes a
CC method that identifies modulators of the target protein, which is a
CC kinesin-like DNA binding protein (known as KID) as cytostatic, cardiantes,
CC immunomodulators and antiinflammatories. Accordingly, through gene
CC therapy, they can be used for the treatment of cancer, hyperplasias,
CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This
CC polypeptide sequence is human KID protein (seqID 2) of the invention.
XX
XX
SQ Sequence 370 AA;

Query Match 100.0%; Score 1878; DB 7; Length 370;
Best Local Similarity 100.0%; Pred. No. 3.8e-182;
Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAAGGSGTQORREMAAASAAISGAGCRCLSKIGATRRPPARVRVAVRLRPFVDGTAGA 60
Db 1 MAAGGSGTQORREMAAASAAISGAGCRCLSKIGATRRPPARVRVAVRLRPFVDGTAGA 60
Qy 61 SDPPCVRGMDSCSLEIANWRNHQETLKYQFDAPYGERSTQDDIYAGSVQPIRLHLLGQN 120
Db 61 SDPPCVRGMDSCSLEIANWRNHQETLKYQFDAPYGERSTQDDIYAGSVQPIRLHLLGQN 120
Qy 121 ASVLAYGPTGAGKTHMLGSPEQPGVPRALMDLLQLTREGAEGRPWALSVTMSYLEIY 180
Db 121 ASVLAYGPTGAGKTHMLGSPEQPGVPRALMDLLQLTREGAEGRPWALSVTMSYLEIY 180
Qy 181 QEKVLDLLDPASGDLVIREDCRGNILIPGLSQPISSFADFPERHFLPASNRVTGATRLN 240
Db 181 QEKVLDLLDPASGDLVIREDCRGNILIPGLSQPISSFADFPERHFLPASNRVTGATRLN 240
Qy 241 QRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGSEDNRRGTNGKLGKESGAINTS 300
Db 241 QRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGSEDNRRGTNGKLGKESGAINTS 300
Qy 301 LFVLGKVVDAALNOGLPRVPYRDSKLTRELQDSLGGSAHSILIANIAPERFYLDTVSALN 360
Db 301 LFVLGKVVDAALNOGLPRVPYRDSKLTRELQDSLGGSAHSILIANIAPERFYLDTVSALN 360
Qy 361 FAARKEVIN 370
Db 361 FAARKEVIN 370

RESULT 2
ADQ60228
ID ADQ60228 standard; protein; 370 AA.
AC ADQ60228;
XX
XX 23-SEP-2004 (first entry)
XX Human microtubule motor protein #1.
XX Human; microtubule motor protein; cellular proliferation disorder;
KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;
KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;
KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.
XX
XX Homo sapiens.
XX US6762043-B1.
XX 13-JUL-2004.
XX
XX 06-MAR-2002; 2002US-00093317.
XX
XX 20-APR-1999; 99US-00295612.
PR 20-JUN-2000; 2000US-00597252.
PR 28-NOV-2000; 2000US-00724224.
XX
XX (CYTO-) CYTOKINETICS INC.

XX Beraud C;
PI
XX
XX WPI; 2004-532491/51.
XX
XX New isolated microtubule motor protein, useful for screening modulators
PT for treating cellular proliferation disorders such as cancer,
PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
PT inflammation.
XX
XX
PS Claim 1; SEQ ID NO 2; 26pp; English.
XX
XX The invention relates to human microtubule motor proteins and the nucleic
CC acids encoding them. The invention also relates to a method of screening
CC for modulators of a motor protein which has microtubule stimulated ATPase
CC activity, a method of testing for ATPase activity of microtubule motor
CC proteins, methods to identify candidate agents that bind to a target
CC protein or act as a modulator of the binding characteristics or
CC biological activity of a target protein, modulators of the target
CC protein, and methods of treating cellular proliferation disorders such as
CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
CC and inflammation, for treating disorders associated with kinesin-like DNA
CC binding protein (KID) and for inhibiting KID. The sequences are used for
CC screening for modulators of motor proteins useful for treating cellular
CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
CC hypertrophy, immune disorders and inflammation, for treating disorders
CC associated with KID and for inhibiting KID and for treating autoimmune
CC diseases, arthritis, graft rejection, inflammatory bowel disease and
CC proliferation induced after medical procedures including surgery and
CC angioplasty. This sequence represents a human microtubule motor protein
CC of the invention. Note: The specification states that this sequence is
CC encoded by the nucleic acid featured as SEQ ID NO:1, but this does not
CC appear to be the case.
XX
SQ Sequence 370 AA;

Query Match 100.0%; Score 1878; DB 8; Length 370;
Best Local Similarity 100.0%; Pred. No. 3.8e-182;
Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAAGGSGTQORREMAAASAAISGAGCRCLSKIGATRRPPARVRVAVRLRPFVDGTAGA 60
Db 1 MAAGGSGTQORREMAAASAAISGAGCRCLSKIGATRRPPARVRVAVRLRPFVDGTAGA 60
Qy 61 SDPPCVRGMDSCSLEIANWRNHQETLKYQFDAPYGERSTQDDIYAGSVQPIRLHLLGQN 120
Db 61 SDPPCVRGMDSCSLEIANWRNHQETLKYQFDAPYGERSTQDDIYAGSVQPIRLHLLGQN 120
Qy 121 ASVLAYGPTGAGKTHMLGSPEQPGVPRALMDLLQLTREGAEGRPWALSVTMSYLEIY 180
Db 121 ASVLAYGPTGAGKTHMLGSPEQPGVPRALMDLLQLTREGAEGRPWALSVTMSYLEIY 180
Qy 181 QEKVLDLLDPASGDLVIREDCRGNILIPGLSQPISSFADFPERHFLPASNRVTGATRLN 240
Db 181 QEKVLDLLDPASGDLVIREDCRGNILIPGLSQPISSFADFPERHFLPASNRVTGATRLN 240
Qy 241 QRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGSEDNRRGTNGKLGKESGAINTS 300
Db 241 QRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGSEDNRRGTNGKLGKESGAINTS 300
Qy 301 LFVLGKVVDAALNOGLPRVPYRDSKLTRELQDSLGGSAHSILIANIAPERFYLDTVSALN 360
Db 301 LFVLGKVVDAALNOGLPRVPYRDSKLTRELQDSLGGSAHSILIANIAPERFYLDTVSALN 360
Qy 361 FAARKEVIN 370
Db 361 FAARKEVIN 370

RESULT 3
ADK40973
ID ADK40973 standard; protein; 490 AA.
XX

AC ADK40973;
 XX 06-MAY-2004 (first entry)
 XX Novel human kinase protein #80.
 XX
 KW cytostatic; immunomodulator; cardiant; neuroprotective; nootropic;
 KW antiparkinsonian; virucide; antibacterial; fungicide; ophthalmological;
 KW analgesic; hypotensive; immunosuppressive; kinase inhibitor; kinase;
 KW cancer; peripheral nervous system; central nervous system;
 KW Alzheimer's disease; Parkinson's disease; multiple sclerosis;
 KW amyotrophic lateral sclerosis; viral infection; prion infection;
 KW ocular disease; migraine; pain; sexual dysfunction; mood disorder;
 KW attention disorder; cognition disorder; hypotension; hypertension;
 KW psychotic disorder; neurological disorder; dyskinesia;
 KW metabolic disorder; organ transplant rejection; enzyme.
 XX
 OS Homo sapiens.
 XX
 XX WO2003057841-A2.
 XX
 XX 17-JUL-2003.
 XX
 XX 31-DEC-2002; 2002WO-US041687.
 XX
 XX 31-DEC-2001; 2001US-0343169P.
 XX
 XX (GRIG/) GRIGORIEV I V.
 XX (SUDA/) SUDARSANAM S.
 XX
 XX Grigoriev IV, Sudarsanam S;
 XX
 XX WPI; 2003-587115/55.
 XX
 XX New isolated, enriched or purified nucleic acid molecule encoding a
 XX kinase polypeptide, useful for treating cancer, immune-related diseases,
 XX cardiovascular disease, brain or neuronal-associated diseases and
 XX metabolic disorders.
 XX
 XX Claim 1; SEQ ID NO 80; 491pp; English.
 XX
 XX The invention relates to novel isolated, enriched or purified nucleic acid
 XX molecules encoding a kinase polypeptide. The nucleic acid molecule
 XX comprises a sequence that: (a) encodes a kinase polypeptide; (b) is a
 XX complement of (a); (c) hybridizes under stringent conditions to (a) and
 XX encodes a naturally occurring kinase polypeptide; (d) encodes the
 XX polypeptide in (a), except that it lacks one or more, but not all, of an
 XX N-terminal domain, C-terminal catalytic domain, a catalytic region, a C-
 XX terminal tail; or (e) is a complement of (d). The nucleic acid molecules,
 XX polypeptides, methods and substance are useful for treating cancers,
 XX immune-related diseases or disorders, cardiovascular disease, brain or
 XX neuronal-associated diseases, and metabolic disorders. The disorders are
 XX preferably cancers of the tissues or of hematopoietic origin, diseases of
 XX the central or peripheral nervous system, Alzheimer's disease,
 XX Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis,
 XX viral infections, infections caused by prions, infections caused by
 XX bacteria, infections caused by fungi, ocular diseases, migraines, pain,
 XX sexual dysfunction, mood disorders, attention disorders, cognition
 XX disorders, hypotension, hypertension, psychotic disorders, neurological
 XX disorders, dyskinesias, metabolic disorders and organ transplant
 XX rejection. This sequence corresponds to one of the kinase polypeptides of
 XX the invention.
 XX
 XX Sequence 490 AA;
 XX
 XX Query Match 100.0%; Score 1878; DB 7; Length 490;
 XX Best Local Similarity 100.0%; Pred. No. 5.9e-182;
 XX Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 MAAGSTQORREMAAASAAISAGRCRLSKIGATRRPPPARVAVRLRPFVDGTAGA 60
 XX 9 MAAGSTQORREMAAASAAISAGRCRLSKIGATRRPPPARVAVRLRPFVDGTAGA 68

QY 61 SDPPCVRGWDSLEIANRNHOETLKYQFDAYGERSTQODIYAGSVQPIRLHLEGGN 120
 DB 69 SDPPCVRGWDSLEIANRNHOETLKYQFDAYGERSTQODIYAGSVQPIRLHLEGGN 128
 QY 121 ASVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTRECAEGRPWALSVTMSYLEIY 180
 DB 129 ASVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTRECAEGRPWALSVTMSYLEIY 188
 QY 181 QEKVLDDLPASGDLVIREDCRGNLIPGLSQKPISSPADFERHFLPASRNTVGTATRLN 240
 DB 189 QEKVLDDLPASGDLVIREDCRGNLIPGLSQKPISSPADFERHFLPASRNTVGTATRLN 248
 QY 241 ORSSSHAVLVKVDQERLAPFRQREKLYLIDLAGSEDNRTGNKGLRKEGAINTS 300
 DB 249 ORSSSHAVLVKVDQERLAPFRQREKLYLIDLAGSEDNRTGNKGLRKEGAINTS 308
 QY 301 LFLVGKVVVDALNQGLPRVPRYRDSKLTLLQDSLGSSAHSILIANIAPERRFYLDVTVALN 360
 DB 309 LFLVGKVVVDALNQGLPRVPRYRDSKLTLLQDSLGSSAHSILIANIAPERRFYLDVTVALN 368
 QY 361 FAARSKEVIN 370
 DB 369 FAARSKEVIN 378

RESULT 4
 ADQ09240
 ID ADQ09240 standard; protein; 665 AA.
 XX
 XX ADQ09240;
 XX
 XX 23-SEP-2004 (first entry)
 XX
 XX Human KNSL4 protein SEQ ID NO:425.
 XX
 KW thanatos-associated protein; THAP; THAP responsive gene; THAP family;
 KW THAP responsive element; angiogenesis; inflammation; metastasis; cancer;
 KW apoptosis; cardiovascular disease; neurodegenerative disease; chemokine;
 KW angiogenic; antiinflammatory; cardiovascular; cytostatic;
 KW neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;
 KW human.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO2004055050-A2.
 XX
 XX 01-JUL-2004.
 XX
 XX 10-DEC-2003; 2003WO-IB006434.
 XX
 XX 10-DEC-2002; 2002US-0432699P.
 XX 03-JUL-2003; 2003US-0485027P.
 XX
 XX (ENDO-) ENDOCUBE SAS.
 XX (CNRS) CNRS CENT NAT RECH SCI.
 XX
 XX Girard J, Amalric F, Roussigne M, Clouaire T;
 XX WPI; 2004-525034/50.
 XX N-PSDB; ADQ09241.
 XX
 XX Modulating expression of a Thanatos (death)-Associated Protein (THAP)
 XX responsive gene for preventing or treating e.g. cancer or inflammation,
 XX comprises modulating the interaction of a THAP polypeptide with a nucleic
 XX acid.
 XX
 XX Example 47; SEQ ID NO 425; 612pp; English.
 XX
 XX The present invention describes a method for modulating the expression of
 XX a thanatos (death)-associated protein (THAP) responsive gene. The method
 XX comprises modulating the interaction of a THAP-family polypeptide or its
 XX biological fragment with a nucleic acid, and so enhancing or repressing

the expression of the THAP responsive gene. Also described: (1) a method of modulating the expression of a gene responsive to a THAP/chemokine complex; (2) a pharmaceutical composition comprising a THAP responsive element in a pharmaceutical carrier; (3) a transcription factor decoy consisting essentially of a THAP responsive element; (4) a cell comprising a transcription factor decoy described above; (5) methods of modulating the interaction between a nucleic acid and a THAP-family polypeptide or its biological fragment, or a nucleic acid and a THAP/chemokine complex; (6) a vector packaging cell line comprising a cell comprising a viral vector which comprises a promoter operably linked to a nucleic acid encoding a THAP-family polypeptide or its biological fragment; (7) a method of constructing a cell which expresses a recombinant THAP-family polypeptide; (8) a method of ameliorating symptoms associated with a condition mediated by a THAP/chemokine complex; (9) methods of identifying a test compound that modulates transcription at a THAP responsive element or that modulates the transport of a chemokine into the nucleus; (10) methods for reducing the symptoms associated with a condition selected from excessive or insufficient angiogenesis, inflammation, metastasis of a cancerous tissue, excessive or insufficient apoptosis, cardiovascular disease and neurodegenerative diseases; symptoms associated with a condition resulting from the activity of a chemokine or a THAP-family polypeptide in an individual; or symptoms associated with transcriptional repression or activation mediated by a THAP-family polypeptide in an individual; (11) a vector comprising a THAP responsive promoter operably linked to a nucleic acid encoding a detectable product; (12) a genetically engineered cell comprising the vector described above or that expresses a THAP-family polypeptide or its biological fragment; (13) an in vitro transcription reaction comprising a nucleic acid comprising a THAP responsive promoter, ribonucleotides and an RNA polymerase; and (14) an isolated mutant THAP-family polypeptide that does not bind to a chemokine. The pharmaceutical composition has antiangiogenic, antiinflammatory, cardiovascular, cytoskeletal, neuroprotective and osteopathic activities, and can be used as a THAP and THAP synthesis modulator. The composition can be used for modulating the expression of a THAP responsive gene. Modulation is useful for reducing symptoms of conditions such as excessive or insufficient angiogenesis, inflammation, metastasis of a cancerous tissue, excessive or insufficient apoptosis, cardiovascular disease or neurodegenerative diseases. The present sequence is used in the exemplification of the present invention.

Query Match 100.0%; Score 1878; DB 8; Length 665;
Best Local Similarity 100.0%; Pred. No. 9,5e-182; Mismatches 0; Indels 0; Gaps 0;
Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MAAGSTQORREMAAASAAISGAGRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGA 60
1 MAAGSTQORREMAAASAAISGAGRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGA 60

61 SDPPCVRGMDSCSLEIANWRNHQETLYQFDAPFYGERSTOODIYAGSVQPIRLHLEGON 120
61 SDPPCVRGMDSCSLEIANWRNHQETLYQFDAPFYGERSTOODIYAGSVQPIRLHLEGON 120

121 ASVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREBAGRGPWALSVTMSYLEIY 180
121 ASVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREBAGRGPWALSVTMSYLEIY 180

181 QEKVLDDLDPASGDIVREDCRGNILIPGLSQKPISSPADFERHFLPASRNRVTGATRLN 240
181 QEKVLDDLDPASGDIVREDCRGNILIPGLSQKPISSPADFERHFLPASRNRVTGATRLN 240

241 QRSRSHAVLLVKVDQRELLAPFRQREKGLYLIDLAGSEDNRRYGNKGLRKGSAINTS 300
241 QRSRSHAVLLVKVDQRELLAPFRQREKGLYLIDLAGSEDNRRYGNKGLRKGSAINTS 300

301 LFLVLGKVDALNQGPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERFYLTVSALN 360
301 LFLVLGKVDALNQGPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERFYLTVSALN 360

361 FAARSKEVIN 370
|||||

Db 361 FAARSKEVIN 370

RESULT 5
ADC23340
ID ADC23340 standard; protein; 512 AA.
XX
AC ADC23340;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human kinesin-like DNA binding protein (KID) (SeqID 4).
XX
KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID;
KW cytoskeletal; cardiant; immunomodulator; antiinflammatory; gene therapy;
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;
KW cardiac hypertrophy; immune disorder; inflammation.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 2 /note= "Encoded by CA"
FT
XX US6387644-B1.
XX PN 14-MAY-2002.
XX PD 28-NOV-2000; 2000US-00724224.
XX PF 20-APR-1999; 99US-00295612.
XX PR 20-JUN-2000; 2000US-00597292.
XX PA (CYTO-) CYTOKINETICS INC.
XX PI Beraud C;
XX WPI; 2003-706919/67.
XX DR N-PSDB; ADC23339.
XX
PT Identifying a candidate agent as modulator of function of a target
PT protein for treating cellular proliferation disorders by adding a
PT candidate agent to a mixture of the target protein that
PT directly/indirectly produces ADP or phosphate.
XX
PS Claim 1; SEQ ID NO 4; 26pp; English.
XX
CC This invention relates to a novel method for high throughput screening
CC systems used to identify compounds for the treatment of cellular
CC proliferation disorders. Specifically, it refers to candidate agents that
CC are capable of modulating the activity of target proteins having motor
CC domains, such that the target protein directly or indirectly produces ADP
CC or phosphate. Furthermore, this activity can be determined using
CC fluorescence or absorbance readouts. The present invention describes a
CC method that identifies modulators of the target protein, which is a
CC kinesin-like DNA binding protein (known as KID) as cytoskeletal, cardiant,
CC immunomodulators and antiinflammatories. Accordingly, through gene
CC therapy, they can be used for the treatment of cancer, hyperplasias,
CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This
CC polypeptide sequence is human KID protein (SeqID 4) of the invention.
XX
SQ Sequence 512 AA;
Query Match 99.7%; Score 1873; DB 7; Length 512;
Best Local Similarity 100.0%; Pred. No. 2e-181;
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAGSTQORREMAAASAAISGAGRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGA 61
Db 3 AAGSTQORREMAAASAAISGAGRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGA 62
QY 62 DPCCVRGMDSCSLEIANWRNHQETLYQFDAPFYGERSTOODIYAGSVQPIRLHLEGON 121
|||||

```
Db 63 DPPCVRGMDSCSLEIANWRNHQETLKQYDFAYGERSTQODIYAGSVQPIRLHLLGQNA 122
QY 122 SVLAYGPTGAGKTHTMLGSPGQVPIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQ 181
Db 123 SVLAYGPTGAGKTHTMLGSPGQVPIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQ 182
QY 182 EKVLDDLPASGDLVIREDCRGNILIPGLSQPKISSFADFERHFLPASNRRTVGATRLNQ 241
Db 183 EKVLDDLPASGDLVIREDCRGNILIPGLSQPKISSFADFERHFLPASNRRTVGATRLNQ 242
QY 242 RSSRSHAVLLVKVDQERLAPRQREGKLYLIDLAGSEDNRTGNKGLKESGAINTSL 301
Db 243 RSSRSHAVLLVKVDQERLAPRQREGKLYLIDLAGSEDNRTGNKGLKESGAINTSL 302
QY 302 FVLGKVVVDALNOGLPRVPRYRDSKLTLLQDSLGSSAHSILIANIAPERFYLDTSALNF 361
Db 303 FVLGKVVVDALNOGLPRVPRYRDSKLTLLQDSLGSSAHSILIANIAPERFYLDTSALNF 362
QY 362 AARSKEVIN 370
Db 363 AARSKEVIN 371

RESULT 6
ADQ60230
ID ADQ60230 standard; protein; 512 AA.
AC ADQ60230;
DT 23-SEP-2004 (first entry)
DE Human microtubule motor protein #2.
KW Human; microtubule motor protein; cellular proliferation disorder;
KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;
KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;
KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.
XX
OS Homo sapiens.
XX
XX US6762043-B1.
XX
XX 13-JUL-2004.
XX
XX 06-MAR-2002; 2002US-00093317.
XX
XX 20-APR-1999; 99US-00295612.
XX
XX 20-JUN-2000; 2000US-00597292.
XX
XX 28-NOV-2000; 2000US-00724224.
XX
XX (CYTO-) CYTOKINETICS INC.
XX
XX Beraud C;
XX
XX WPI; 2004-532491/51.
XX
XX New isolated microtubule motor protein, useful for screening modulators
XX PT for treating cellular proliferation disorders such as cancer,
XX PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
XX PT inflammation.
XX
XX Claim 1; SEQ ID NO 4; 26pp; English.
XX
XX The invention relates to human microtubule motor proteins and the nucleic
XX CC acids encoding them. The invention also relates to a method of screening
XX CC for modulators of a motor protein which has microtubule stimulated ATPase
XX CC activity, a method of testing for ATPase activity of microtubule motor
XX CC proteins, methods to identify candidate agents that bind to a target
XX CC protein or act as a modulator of the binding characteristics or
XX CC biological activity of a target protein, modulators of the target
XX CC protein, and methods of treating cellular proliferation disorders such as
XX CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
XX CC and inflammation, for treating disorders associated with kinesin-like DNA
```

```
CC binding protein (KID) and for inhibiting KID. The sequences are used for
CC screening for modulators of motor proteins useful for treating cellular
CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
CC hypertrophy, immune disorders and inflammation, for treating disorders
CC associated with KID and for inhibiting KID and for treating autoimmune
CC diseases, arthritis, graft rejection, inflammatory bowel disease and
CC proliferation induced after medical procedures including surgery and
CC angioplasty. This sequence represents a human microtubule motor protein
CC of the invention. Note: The specification states that this sequence is
CC encoded by the nucleic acid featured as SEQ ID NO:3, but this does not
CC appear to be the case.
XX
SQ Sequence 512 AA;
Query Match 99.7%; Score 1873; DB 8; Length 512;
Best Local Similarity 100.0%; Pred. No. 2e-181;
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 AAGGSTQQRREMAAASAAASGACRCLSKIGATRRPPPARVRVAVRLRPFVDGTAGAS 61
Db 3 AAGGSTQQRREMAAASAAASGACRCLSKIGATRRPPPARVRVAVRLRPFVDGTAGAS 62
QY 62 DPPCVRGMDSCSLEIANWRNHQETLKQYDFAYGERSTQODIYAGSVQPIRLHLLGQNA 121
Db 63 DPPCVRGMDSCSLEIANWRNHQETLKQYDFAYGERSTQODIYAGSVQPIRLHLLGQNA 122
QY 122 SVLAYGPTGAGKTHTMLGSPGQVPIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQ 181
Db 123 SVLAYGPTGAGKTHTMLGSPGQVPIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQ 182
QY 182 EKVLDDLPASGDLVIREDCRGNILIPGLSQPKISSFADFERHFLPASNRRTVGATRLNQ 241
Db 183 EKVLDDLPASGDLVIREDCRGNILIPGLSQPKISSFADFERHFLPASNRRTVGATRLNQ 242
QY 242 RSSRSHAVLLVKVDQERLAPRQREGKLYLIDLAGSEDNRTGNKGLKESGAINTSL 301
Db 243 RSSRSHAVLLVKVDQERLAPRQREGKLYLIDLAGSEDNRTGNKGLKESGAINTSL 302
QY 302 FVLGKVVVDALNOGLPRVPRYRDSKLTLLQDSLGSSAHSILIANIAPERFYLDTSALNF 361
Db 303 FVLGKVVVDALNOGLPRVPRYRDSKLTLLQDSLGSSAHSILIANIAPERFYLDTSALNF 362
QY 362 AARSKEVIN 370
Db 363 AARSKEVIN 371

RESULT 7
ADC23342
ID ADC23342 standard; protein; 346 AA.
AC ADC23342;
XX
XX 18-DEC-2003 (first entry)
XX
XX Human kinesin-like DNA binding protein (KID) (SeqID 6).
XX
XX human; enzyme; motor domain; kinesin-like DNA binding protein; KID;
XX KW cytoskeletal; cardiac; immunomodulator; antiinflammatory; gene therapy;
XX KW cancer; hyperplasia; restenosis; cellular proliferation disorder;
XX KW cardiac hypertrophy; immune disorder; inflammation.
XX
XX OS Homo sapiens.
XX
XX US6387644-B1.
XX
XX 14-MAY-2002.
XX
XX 28-NOV-2000; 2000US-00724224.
XX
XX 20-APR-1999; 99US-00295612.
XX
XX 20-JUN-2000; 2000US-00597292.
XX
```

PA (CYTO-) CYTOKINETICS INC.
 XX Beraud C;
 PI WPI; 2003-706919/67.
 XX N-PSDB; ADC23341.
 DR
 DR
 DR
 XX Identifying a candidate agent as modulator of function of a target
 PT protein for treating cellular proliferation disorders by adding a
 PT candidate agent to a mixture of the target protein that
 PT directly/indirectly produces ADP or phosphate.
 XX
 XX Claim 1; SEQ ID NO 6; 26pp; English.
 PS
 XX This invention relates to a novel method for high throughput screening
 CC systems used to identify compounds for the treatment of cellular
 CC proliferation disorders. Specifically, it refers to candidate agents that
 CC are capable of modulating the activity of target proteins having motor
 CC domains, such that the target protein directly or indirectly produces ADP
 CC or phosphate. Furthermore, this activity can be determined using
 CC fluorescence or absorbance readouts. The present invention describes a
 CC method that identifies modulators of the target protein, which is a
 CC kinesin-like DNA binding protein (known as KID) as cytostatic, cardants,
 CC immunomodulators and antiinflammatory. Accordingly, through gene
 CC therapy, they can be used for the treatment of cancer, hyperplasias,
 CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This
 CC polypeptide sequence is human KID protein (SeqID 6) of the invention.
 XX
 XX Sequence 346 AA;
 SQ
 Query Match 93.9%; Score 1763; DB 7; Length 346;
 Best Local Similarity 100.0%; Pred. No. 1.8e-170;
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 26 GRCLSKIGATRRPPPPARVRVAVLRPFVDGTAGASDPPCVRGWMDSCSLEIANWRNHQET 85
 DB 2 GRCLSKIGATRRPPPPARVRVAVLRPFVDGTAGASDPPCVRGWMDSCSLEIANWRNHQET 61
 QY 86 LKQFDIFYGERSQDIIYAGSVQPIRLHLEGGQNASVLAAYGPTGAGKTHMLGSPQPG 145
 DB 62 LKQFDIFYGERSQDIIYAGSVQPIRLHLEGGQNASVLAAYGPTGAGKTHMLGSPQPG 121
 QY 146 VIPRALMDLLQTLTREGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDVIVREDCRGNI 205
 DB 122 VIPRALMDLLQTLTREGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDVIVREDCRGNI 181
 QY 206 LIPGLSQKPISSFADFERHFLPASRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 265
 DB 182 LIPGLSQKPISSFADFERHFLPASRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 241
 QY 266 REGKYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVGLKVVDALNOGLPRVPYRDSKL 325
 DB 242 REGKYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVGLKVVDALNOGLPRVPYRDSKL 301
 QY 326 TRLLQDSLGSGSAHSILIANIAPERRFYDITVSALNFAARSKEVIN 370
 DB 302 TRLLQDSLGSGSAHSILIANIAPERRFYDITVSALNFAARSKEVIN 346
 RESULT 8
 ADQ60232
 ID ADQ60232 standard; protein; 346 AA.
 XX
 AC ADQ60232;
 XX
 DT 23-SEP-2004 (first entry)
 XX
 DE Human microtubule motor protein #3.
 XX
 KW Human; microtubule motor protein; cellular proliferation disorder;
 KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;
 KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;
 KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.

XX Homo sapiens.
 OS
 XX US6762043-B1.
 PN
 XX
 XX 13-JUL-2004.
 PD
 XX
 XX 06-MAR-2002; 2002US-00093317.
 PF
 XX 20-APR-1999; 99US-00295612.
 XX
 PR 20-JUN-2000; 2000US-00597292.
 PR
 PR 28-NOV-2000; 2000US-00724224.
 PR
 XX (CYTO-) CYTOKINETICS INC.
 PA
 XX Beraud C;
 XX
 PI WPI; 2004-532491/51.
 XX
 DR N-PSDB; ADQ60231.
 DR
 DR New isolated microtubule motor protein, useful for screening modulators
 XX for treating cellular proliferation disorders such as cancer,
 PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
 PT inflammation.
 PT
 XX Claim 1; SEQ ID NO 6; 26pp; English.
 PS
 XX The invention relates to human microtubule motor proteins and the nucleic
 CC acids encoding them. The invention also relates to a method of screening
 CC for modulators of a motor protein which has microtubule stimulated ATPase
 CC activity, a method of testing for ATPase activity of microtubule motor
 CC proteins, methods to identify candidate agents that bind to a target
 CC protein or act as a modulator of the binding characteristics or
 CC biological activity of a target protein, modulators of the target
 CC protein, and methods of treating cellular proliferation disorders such as
 CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
 CC and inflammation, for treating disorders associated with kinesin-like DNA
 CC binding protein (KID) and for inhibiting KID. The sequences are used for
 CC screening for modulators of motor proteins useful for treating cellular
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
 CC hypertrophy, immune disorders and inflammation, for treating disorders
 CC associated with KID and for inhibiting KID and for treating autoimmune
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and
 CC proliferation induced after medical procedures including surgery and
 CC angioplasty. This sequence represents a human microtubule motor protein
 CC of the invention.
 XX
 SQ Sequence 346 AA;
 Query Match 93.9%; Score 1763; DB 8; Length 346;
 Best Local Similarity 100.0%; Pred. No. 1.8e-170;
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 26 GRCLSKIGATRRPPPPARVRVAVLRPFVDGTAGASDPPCVRGWMDSCSLEIANWRNHQET 85
 DB 2 GRCLSKIGATRRPPPPARVRVAVLRPFVDGTAGASDPPCVRGWMDSCSLEIANWRNHQET 61
 QY 86 LKQFDIFYGERSQDIIYAGSVQPIRLHLEGGQNASVLAAYGPTGAGKTHMLGSPQPG 145
 DB 62 LKQFDIFYGERSQDIIYAGSVQPIRLHLEGGQNASVLAAYGPTGAGKTHMLGSPQPG 121
 QY 146 VIPRALMDLLQTLTREGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDVIVREDCRGNI 205
 DB 122 VIPRALMDLLQTLTREGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDVIVREDCRGNI 181
 QY 206 LIPGLSQKPISSFADFERHFLPASRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 265
 DB 182 LIPGLSQKPISSFADFERHFLPASRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 241
 QY 266 REGKYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVGLKVVDALNOGLPRVPYRDSKL 325
 DB 242 REGKYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVGLKVVDALNOGLPRVPYRDSKL 301

QY 326 TRLLQSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 370
|||||
Db 302 TRLLQSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
|||||

RESULT 9

AAB56650
ID AAB56650 standard; protein; 460 AA.

AC AAB56650;

DT 13-MAR-2001 (first entry)

DE Human prostate cancer antigen protein sequence SEQ ID NO:1228.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease.

OS Homo sapiens.

XX WO200055174-A1.

PN 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US005988.

XX 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.

XX Rosen CA, Ruben SM;

PI WPI; 2000-587513/55.

DR N-PSDB; AAF15853.

XX Prostate cancer associated gene sequences, referred to as prostate cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as prostate cancer.

XX Claim 11; Page 1649-1651; 2338pp; English.

XX AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention

XX Sequence 460 AA;

Query Match 93.9%; Score 1763; DB 3; Length 460;
Best Local Similarity 99.4%; Pred. No. 2.8e-170;
Matches 345; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 24 GAGCRLSKIGATRRPPPARVAVLRPFVDTAGADPPCVRGWDCSLFIANWRNHQ 83
|||||

Db 2 GAGCRLSKIGATRRPPPARVAVLRPFVDTAGADPPCVRGWDCSLFIANWRNHQ 61
|||||

QY 84 EFLKYQDAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPEQ 143
|||||

Db 62 ETLKYQDAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPEQ 121
|||||
QY 144 PGVIPRALMDLLQLTREEGAGRPWALSVTMSVLEIYQEKVLDLLDPASGDLVIREDCRG 203
|||||
Db 122 PGVIPRALMDLLQLTREEGAGRPWALSVTMSVLEIYQEKVLDLLDPASGDLVIREDCRG 181
|||||
QY 204 NILIPGLSQKPISSPADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQERLAPF 263
|||||
Db 182 NILIPGLSQKPISSPADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQERLAPF 241
|||||
QY 264 ROREGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVGLKVVDALNOGLPRVYRDS 323
|||||
Db 242 ROREGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVGLKVVDALNOGLPRVYRDS 301
|||||
QY 324 KLTRLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 370
|||||
Db 302 KLTRLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 348
|||||

RESULT 10

ADC23344

ID ADC23344 standard; protein; 487 AA.

XX ADC23344;

AC ADC23344;

DT 18-DEC-2003 (first entry)

XX Human kinesin-like DNA binding protein (KID) (SeqID 8).

XX human; enzyme; motor domain; kinesin-like DNA binding protein; KID;
KW cytostatic; cardiac; immunomodulator; antiinflammatory; gene therapy;
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;
KW cardiac hypertrophy; immune disorder; inflammation.

OS Homo sapiens.

XX US6387644-B1.

PN 14-MAY-2002.

XX 28-NOV-2000; 2000US-00724224.

XX 20-APR-1999; 99US-00295612.

XX 20-JUN-2000; 2000US-00597292.

XX (CYTO-) CYTOKINETICS INC.

XX Beraud C;

XX WPI; 2003-706919/67.

XX N-PSDB; ADC23343.

XX Identifying a candidate agent as modulator of function of a target
PT protein for treating cellular proliferation disorders by adding a
PT candidate agent to a mixture of the target protein that
PT directly/indirectly produces ADP or phosphate.

XX Claim 1; SEQ ID NO 8; 26pp; English.

XX This invention relates to a novel method for high throughput screening
CC systems used to identify compounds for the treatment of cellular
CC proliferation disorders. Specifically, it refers to candidate agents that
CC are capable of modulating the activity of target proteins having motor
CC domains, such that the target protein directly or indirectly produces ADP
CC or phosphate. Furthermore, this activity can be determined using
CC fluorescence or absorbance readouts. The present invention describes a
CC method that identifies modulators of the target protein, which is a
CC kinesin-like DNA binding protein (known as KID) as cytostatic, cardiacs,
CC immunomodulators and antiinflammatories. Accordingly, through gene
CC therapy, they can be used for the treatment of cancer, hyperplasia,
CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This
CC polypeptide sequence is human KID protein (SeqID 8) of the invention.

SQ	Sequence 487 AA;	
	Query Match 93.9%; Score 1763; DB 7; Length 487;	
	Best Local Similarity 100.0%; Pred. No. 3.1e-170;	
	Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	26 GRCLSKIGATRRPPPARVRVAVLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 85	
DB	2 GRCLSKIGATRRPPPARVRVAVLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61	
QY	86 LKYQDFAFYGERSTQODIYAGSVQPIRLHLEQONASVLAAYGTGAGKTHTMLGSPQPG 145	
DB	62 LKYQDFAFYGERSTQODIYAGSVQPIRLHLEQONASVLAAYGTGAGKTHTMLGSPQPG 121	
QY	146 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQEKVLDLDDPASGDLVIREDCRGN 205	
DB	122 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQEKVLDLDDPASGDLVIREDCRGN 181	
QY	206 LIPGLSQKPISSFADFERRHFLPASRRNTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 265	
DB	182 LIPGLSQKPISSFADFERRHFLPASRRNTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 241	
QY	266 REGKYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVGLKVVDALNOGLPRVPYRDSKL 325	
DB	242 REGKYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVGLKVVDALNOGLPRVPYRDSKL 301	
QY	326 TRLLQDSLGGSAHSILIANIAPERFYLDTVSALNFAARKEVIN 370	
DB	302 TRLLQDSLGGSAHSILIANIAPERFYLDTVSALNFAARKEVIN 346	
RESULT 11		
ID	ADQ60234	
XX	ADQ60234 standard; protein; 487 AA.	
AC	ADQ60234;	
DT	23-SEP-2004 (first entry)	
DE	Human microtubule motor protein #4.	
XX	Human; microtubule motor protein; cellular proliferation disorder;	
KW	cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;	
KW	inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;	
KW	arthritis; graft rejection; inflammatory bowel disease; angioplasty.	
OS	Homo sapiens.	
XX	US6762043-B1.	
PN	13-JUL-2004.	
XX	06-MAR-2002; 2002US-00093317.	
XX	20-APR-1999; 99US-00295612.	
PR	20-JUN-2000; 2000US-00597292.	
PR	28-NOV-2000; 2000US-00724224.	
XX	(CYTO-) CYTOKINETICS INC.	
PA	Beraud C;	
XX	WPI; 2004-532491/51.	
DR	N-PSDB; ADQ60233.	
XX	New isolated microtubule motor protein, useful for screening modulators	
PT	for treating cellular proliferation disorders such as cancer,	
PT	hyperplasias, restenosis, cardiac hypertrophy, immune disorders and	
PT	inflammation.	
XX	Claim 1; SEQ ID NO 8; 26pp; English.	
PS	The invention relates to human microtubule motor proteins and the nucleic	
XX		
CC		

CC	acids encoding them. The invention also relates to a method of screening	
CC	for modulators of a motor protein which has microtubule stimulated ATPase	
CC	activity, a method of testing for ATPase activity of microtubule motor	
CC	proteins, methods to identify candidate agents that bind to a target	
CC	protein or act as a modulator of the binding characteristics or	
CC	biological activity of a target protein, modulators of the target	
CC	protein, and methods of treating cellular proliferation disorders such as	
CC	cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders	
CC	and inflammation, for treating disorders associated with kinesin-like DNA	
CC	binding protein (KID) and for inhibiting KID. The sequences are used for	
CC	screening for modulators of motor proteins useful for treating cellular	
CC	proliferation disorders such as cancer, hyperplasias, restenosis, cardiac	
CC	hypertrophy, immune disorders and inflammation, for treating disorders	
CC	associated with KID and for inhibiting KID and for treating autoimmune	
CC	diseases, arthritis, graft rejection, inflammatory bowel disease and	
CC	proliferation induced after medical procedures including surgery and	
CC	angioplasty. This sequence represents a human microtubule motor protein	
CC	of the invention.	
XX	Sequence 487 AA;	
SQ	Query Match 93.9%; Score 1763; DB 8; Length 487;	
	Best Local Similarity 100.0%; Pred. No. 3.1e-170;	
	Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	26 GRCLSKIGATRRPPPARVRVAVLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 85	
DB	2 GRCLSKIGATRRPPPARVRVAVLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61	
QY	86 LKYQDFAFYGERSTQODIYAGSVQPIRLHLEQONASVLAAYGTGAGKTHTMLGSPQPG 145	
DB	62 LKYQDFAFYGERSTQODIYAGSVQPIRLHLEQONASVLAAYGTGAGKTHTMLGSPQPG 121	
QY	146 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQEKVLDLDDPASGDLVIREDCRGN 205	
DB	122 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQEKVLDLDDPASGDLVIREDCRGN 181	
QY	206 LIPGLSQKPISSFADFERRHFLPASRRNTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 265	
DB	182 LIPGLSQKPISSFADFERRHFLPASRRNTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 241	
QY	266 REGKYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVGLKVVDALNOGLPRVPYRDSKL 325	
DB	242 REGKYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVGLKVVDALNOGLPRVPYRDSKL 301	
QY	326 TRLLQDSLGGSAHSILIANIAPERFYLDTVSALNFAARKEVIN 370	
DB	302 TRLLQDSLGGSAHSILIANIAPERFYLDTVSALNFAARKEVIN 346	
RESULT 12		
ABB71112		
ID	ABB71112 standard; protein; 784 AA.	
XX	ABB71112;	
AC	ABB71112;	
XX	26-MAR-2002 (first entry)	
DT		
XX	Drosophila melanogaster polypeptide SEQ ID NO 40128.	
DE	Drosophila; developmental biology; cell signalling; insecticide;	
XX	pharmaceutical.	
KW	Drosophila melanogaster.	
OS	WO200171042-A2.	
XX	27-SEP-2001.	
PN	23-MAR-2001; 2001WO-US009231.	
XX	23-MAR-2000; 2000US-0191637P.	
PR	11-JUL-2000; 2000US-00614150.	

cell proliferative disorder; cancer; atherosclerosis; viral infection; HIV; neurological disorder; Parkinson's disease; Alzheimer's disease; stroke; human.

Homo sapiens.

Key	Location/Qualifiers
-----	---------------------

FT	Domain	9	/note = kinesin motor catalytic domain
FT			
FT	Domain	11..25	/note = kinesin motor domain
FT			
FT	Modified-site	22	/note = "potential phosphorylation site"
FT			
FT	Domain	62..77	/note = kinesin motor domain
FT			
FT	Domain	63..647	/note = kinesin motor domain
FT			
FT	Modified-site	65	/note = "potential phosphorylation site"
FT			
FT	Modified-site	84	/note = "potential phosphorylation site"
FT			
FT	Modified-site	87	/note = "potential phosphorylation site"
FT			
FT	Domain	95..116	/note = kinesin motor domain
FT			
FT	Modified-site	97	/note = "potential glycosylation site"
FT			
FT	Modified-site	117	/note = "potential phosphorylation site"
FT			
FT	Domain	122..135	/note = kinesin motor domain
FT			
FT	Modified-site	127	/note = "potential phosphorylation site"
FT			
FT	Modified-site	138	/note = "potential phosphorylation site"
FT			
FT	Domain	144..162	/note = kinesin motor domain
FT			
FT	Modified-site	149	/note = "potential phosphorylation site"
FT			
FT	Domain	202..226	/note = kinesin motor domain
FT			
FT	Modified-site	214	/note = "potential glycosylation site"
FT			
FT	Modified-site	217	/note = "potential phosphorylation site"
FT			
FT	Modified-site	227	/note = "potential phosphorylation site"
FT			
FT	Domain	248..287	/note = kinesin motor domain
FT			
FT	Modified-site	254	/note = "potential phosphorylation site"
FT			
FT	Modified-site	274	/note = "potential glycosylation site"
FT			
FT	Modified-site	288	/note = "potential phosphorylation site"
FT			
FT	Modified-site	292	/note = "potential phosphorylation site"
FT			
FT	Domain	295..325	/note = kinesin motor domain
FT			
FT	Modified-site	328	/note = "potential phosphorylation site"
FT			
FT	Modified-site	356	/note = "potential glycosylation site"
FT			
FT	Modified-site	386	/note = "potential phosphorylation site"
FT			
FT	Region	397..399	/note = "leucine zipper pattern"
FT			
FT	Modified-site	454	/note = "potential phosphorylation site"
FT			
FT	Modified-site	469	/note = "potential phosphorylation site"
FT			
FT	Modified-site	489	/note = "potential phosphorylation site"
FT			

FT	Modified-site	495	"potential phosphorylation site"
FT		/note=	"potential phosphorylation site"
FT	Region	518..539	
FT		/note=	"leucine zipper pattern"
FT	Region	525..546	
FT		/note=	"leucine zipper pattern"
FT	Modified-site	605	"potential phosphorylation site"
FT		/note=	"potential phosphorylation site"
FT	Modified-site	649	
FT		/note=	"potential phosphorylation site"
FT	Modified-site	678	
FT		/note=	"potential phosphorylation site"
FT	Modified-site	682	
FT		/note=	"potential phosphorylation site"
FT	Modified-site	703	
FT		/note=	"potential phosphorylation site"
FT	Modified-site	753	
FT		/note=	"potential phosphorylation site"
FT	Modified-site	760	
FT		/note=	"potential glycosylation site"
FT	Modified-site	767	
FT		/note=	"potential phosphorylation site"
FT	Modified-site	790	
FT		/note=	"potential phosphorylation site"
FT	Modified-site	805	
FT		/note=	"potential phosphorylation site"
FT	Modified-site	814	
FT		/note=	"potential phosphorylation site"
FT	Modified-site	816	
FT		/note=	"potential phosphorylation site"
FT	Modified-site	845	
FT		/note=	"potential glycosylation site"
FT	Modified-site	849	
FT		/note=	"potential phosphorylation site"
FT	Modified-site	893	
FT		/note=	"potential phosphorylation site"
FT	Modified-site	973	
FT		/note=	"potential phosphorylation site"
FT	Modified-site	977	
FT		/note=	"potential phosphorylation site"
XX			
XX	WO2004029205-A2..		
XX			
XX	08-APR-2004..		
XX			
XX	24-SEP-2003; 2003WO-US030198..		
XX			
XX	27-SEP-2002; 2002US-0414227P..		
PR	18-NOV-2002; 2002US-0427594P..		
PR	07-FEB-2003; 2003US-0445724P..		
PR	07-MAR-2003; 2003US-0453277P..		
XX			
XX	(INCY-) INCYTE CORP..		
PA			
XX	Swarnakar A, Ho A, Hafalia AJA, Emerling BM, Ison C		
PI	Chien D, Lu DAM, Yang YG, Yue H, Murage J, Griffin		
PI	Wang JT, Marquis JP, Blake JJ, Chawla NK, Lee S, B		
PI	Richardson TW, Bhatia UG, Zheng W, Sprague WW, Tang		
XX			
XX	WPI; 2004-30515/28..		
DR	N-PSDB; ADO44181..		
DR			
XX	New human structural and cytoskeleton-associated proteol		
PT	polynucleotides for diagnosing, preventing or treating		
PT	associated with aberrant protein expression, e.g. cancer		
PT	atherosclerosis, HIV or stroke.		
XX			
XX	Claim 1; Page 150-152; 205pp; English.		
PS			
XX			
CC	The present sequence represents a structural and cytoske		
CC	polypeptide (SCAP). The SCAP polypeptides and polynucle		
CC	invention are useful for diagnosing, preventing or treati		
CC	conditions associated with aberrant expression of SCAP.		

Db 304 ---HIPYNSKUTRLLKDSLGNCQTIMIAAVSPSSVFYDDTYNTLKYANRAKDI 355

Search completed: November 5, 2004, 18:44:37
Job time : 78.9417 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:44:52 ; Search time 86.8758 Seconds
(without alignments)
2081.726 Million cell updates/sec

Title: US-10-797-893-4

Perfect score: 2589

Sequence: 1 MPAAGSTQQRREMAAASA.....LEAKMLAQAKAEKENHCPTM 512

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2589	100.0	512	16	US-10-797-893-4
2	2467	95.3	487	16	US-10-797-893-8
3	2419	93.4	490	15	US-10-334-143-80
4	2151	83.1	460	9	US-09-925-300-1228
5	1873	72.3	370	16	US-10-797-893-2
6	1763	68.1	346	16	US-10-797-893-6
7	615.5	23.8	589	17	US-10-425-115-320209
8	601	23.2	377	16	US-10-437-963-159031
9	579	22.4	776	16	US-10-437-963-103943
10	578.5	22.3	854	17	US-10-425-115-332919
11	578.5	22.3	854	15	US-10-425-114-59708
12	578.5	22.3	1237	15	US-10-334-143-33
13	574.5	22.2	1232	14	US-10-116-712-670

14	574.5	22.2	1232	16	US-10-408-765A-2153	Sequence 2153, Ap
15	574.5	22.2	1235	15	US-10-334-143-8	Sequence 8, Appli
16	569	22.0	1030	15	US-10-425-114-62748	Sequence 62748, A
17	566	21.9	1001	17	US-10-425-115-231231	Sequence 231231,
18	564.5	21.8	548	15	US-10-108-260A-2692	Sequence 2692, Ap
19	561	21.7	1388	14	US-10-146-473-82	Sequence 82, Appl
20	561	21.7	1388	15	US-10-173-999-32	Sequence 32, Appl
21	561	21.7	1388	15	US-10-332-089-2	Sequence 2, Appli
22	561	21.7	1388	15	US-10-188-832-164	Sequence 164, App
23	560	21.6	992	17	US-10-739-930-7903	Sequence 7903, Ap
24	559.5	21.6	1401	15	US-10-287-226-142	Sequence 142, App
25	557.5	21.5	672	16	US-10-408-765A-1664	Sequence 1664, Ap
26	557.5	21.5	1232	14	US-10-116-712-664	Sequence 664, App
27	557.5	21.5	1232	14	US-10-116-712-669	Sequence 669, App
28	557	21.5	1324	15	US-10-287-226-314	Sequence 314, App
29	552.5	21.3	757	15	US-10-220-120-366	Sequence 366, App
30	552.5	21.3	757	16	US-10-363-829-316	Sequence 316, App
31	550	21.2	864	9	US-09-883-096-2	Sequence 1, Appli
32	546.5	21.1	1103	9	US-09-847-874A-1	Sequence 1, Appli
33	546.5	21.1	1103	14	US-10-458-162-1	Sequence 1477, Ap
34	546.5	21.1	1826	16	US-10-408-765A-1477	Sequence 306, App
35	545.5	21.1	834	15	US-10-287-226-306	Sequence 25, Appl
36	544.5	21.0	935	14	US-10-080-608A-25	Sequence 114, App
37	544.5	21.0	935	14	US-10-332-089-6	Sequence 6, Appli
38	543	21.0	409	15	US-10-370-685-114	Sequence 23, Appl
39	543	21.0	928	14	US-10-080-608A-23	Sequence 112, App
40	543	21.0	928	14	US-10-370-685-112	Sequence 176714,
41	541.5	20.9	1382	16	US-10-437-963-176714	Sequence 6, Appli
42	535	20.7	513	17	US-10-601-036-6	Sequence 2, Appli
43	535	20.6	1057	17	US-10-601-036-2	Sequence 472, App
44	532	20.5	1056	14	US-10-282-174-472	Sequence 474, App
45	532	20.5	1056	14	US-10-282-174-474	

ALIGNMENTS

RESULT 1

US-10-797-893-4
; Sequence 4, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10797,893
; PRIOR FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Human
US-10-797-893-4

Query Match	100.0%	Score 2589;	DB 16;	Length 512;
Best Local Similarity	100.0%	Pred. No. 3.3e-209;		
Matches 512;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MPAAGSTQQRREMAAASAAAI	SGAGRCRLSKIGATRRPPPARVVRVAVRLRPVVDGTAG	60
Db	1	MPAAGSTQQRREMAAASAAAI	SGAGRCRLSKIGATRRPPPARVVRVAVRLRPVVDGTAG	60
Qy	61	ASDPPCVRGMDSCSLLEIANWRNHQETLKYQFDAPYGERSTQDDIYAGSVQPIRLHLEGG	120	
Db	61	ASDPPCVRGMDSCSLLEIANWRNHQETLKYQFDAPYGERSTQDDIYAGSVQPIRLHLEGG	120	
Qy	121	NASVLAVGPTGAGKTHMLGSPQGVIPRALMDLLQLTREEGAEGRPWALSVTMSYLEI	180	

```
Db 121 NASVLAAYGPTGAGKTHMLGSPQPGVPRALMDLLQLTREAGRPNWALSVMYSYLEI 180
QY 181 YQKVLDDLDPAGDVLVIREDCRGNLILPGLSQKPISSFADFERHFLPASRNRVTGATRL 240
Db 181 YQKVLDDLDPAGDVLVIREDCRGNLILPGLSQKPISSFADFERHFLPASRNRVTGATRL 240
QY 241 NQSSRSHAVLLVKVDQERLAPFQREGKLYLIDLAGSEDNRRTGNKGLRKESGAIN 300
Db 241 NQSSRSHAVLLVKVDQERLAPFQREGKLYLIDLAGSEDNRRTGNKGLRKESGAIN 300
QY 301 SLFVLGKVVADALNOGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDVT 360
Db 301 SLFVLGKVVADALNOGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDVT 360
QY 361 NFAARKEVINRPTNESLOPHALGPVKLSQKELLGPPAKARGPEEBEIGSPPEMAAP 420
Db 361 NFAARKEVINRPTNESLOPHALGPVKLSQKELLGPPAKARGPEEBEIGSPPEMAAP 420
QY 421 ASASOKLSPLQKLSMDPAMLERLLSLDRLLASQSGAPLLSTPKRERWMLMKTVEED 480
Db 421 ASASOKLSPLQKLSMDPAMLERLLSLDRLLASQSGAPLLSTPKRERWMLMKTVEED 480
QY 481 LEIERLTKQKELEAKMLAQKAEKENHCP 512
Db 481 LEIERLTKQKELEAKMLAQKAEKENHCP 512

RESULT 2
US-10-797-893-8
; Sequence 8, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/797,893
; PRIOR FILING DATE: 2004-03-09
; PRIOR FILING DATE: 2000-11-28
; PRIOR FILING DATE: 2000-11-28
; PRIOR FILING DATE: 09/597,292
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Human
US-10-797-893-8

Query Match 95.3%; Score 2467; DB 16; Length 487;
Best Local Similarity 100.0%; Pred. No. 5.9e-199;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 GRCRLSKIGATRRPPPARVAVRLPFPVDTAGASDPPCVRGMDSCSLEIANRNHQT 86
Db 2 GRCRLSKIGATRRPPPARVAVRLPFPVDTAGASDPPCVRGMDSCSLEIANRNHQT 61
QY 87 LKYPDAFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAYGPTGAGKTHMLGSPQPG 146
Db 62 LKYPDAFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAYGPTGAGKTHMLGSPQPG 121
QY 147 VTPRALMDLLQTLTREAGRPNWALSVMYSYLEIYQKVLDDLDPAGDVLVIREDCR 206
Db 122 VTPRALMDLLQTLTREAGRPNWALSVMYSYLEIYQKVLDDLDPAGDVLVIREDCR 181
QY 207 LIPGLSQKPISSFADFERHFLPASRNRVTGATRLNQRSSRSHAVLLVKVDQERLAP 266
Db 182 LIPGLSQKPISSFADFERHFLPASRNRVTGATRLNQRSSRSHAVLLVKVDQERLAP 241
QY 267 REGKLYLIDLAGSEDNRRTGNKGLRKESCAINTSLFVLGKVVADALNOGLPRVPYR 326
```

```
Db 242 REGKLYLIDLAGSEDNRRTGNKGLRKESCAINTSLFVLGKVVADALNOGLPRVPYR 301
QY 327 TRLQDSLGSAHSILIANIAPERRFYLDVTGALNFAARKEVINRPTNESLOPHALGP 386
Db 302 TRLQDSLGSAHSILIANIAPERRFYLDVTGALNFAARKEVINRPTNESLOPHALGP 361
QY 387 VKLSQKELLGPPAKARGPEEBEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 446
Db 362 VKLSQKELLGPPAKARGPEEBEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 421
QY 447 LDRLLASQSGAPLLSTPKRERWMLMKTVEEDLEIERLTKQKELEAKMLAQKAEKE 506
Db 422 LDRLLASQSGAPLLSTPKRERWMLMKTVEEDLEIERLTKQKELEAKMLAQKAEKE 481
QY 507 NHCPTM 512
Db 482 NHCPTM 487

RESULT 3
US-10-334-143-80
; Sequence 80, Application US/10334143
; Publication No. US20040009549A1
; GENERAL INFORMATION:
; APPLICANT: GRIGORIEV, IGOR VVACHESLAVOVICH
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
; FILE REFERENCE: 038602/1543
; CURRENT APPLICATION NUMBER: US/10/334,143
; PRIOR FILING DATE: 2002-12-31
; PRIOR FILING DATE: 60/343,169
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 80
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-334-143-80

Query Match 93.4%; Score 2419; DB 15; Length 490;
Best Local Similarity 100.0%; Pred. No. 6.6e-195;
Matches 479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAGGSTQQRREMAAASAAAISGAGRCRLSKIGATRRPPPARVAVRLPFPVDTAGAS 62
Db 10 AAGGSTQQRREMAAASAAAISGAGRCRLSKIGATRRPPPARVAVRLPFPVDTAGAS 69
QY 63 DPPCVRGMDSCSLEIANRNHQTLYKQYDAPFVGERSTQODIYAGSVQPIRLHLLLEG 122
Db 70 DPPCVRGMDSCSLEIANRNHQTLYKQYDAPFVGERSTQODIYAGSVQPIRLHLLLEG 129
QY 123 SVLAYGPTGAGKTHMLGSPQPGVTPRALMDLLQTLTREAGRPNWALSVMYSYLEIYQ 182
Db 130 SVLAYGPTGAGKTHMLGSPQPGVTPRALMDLLQTLTREAGRPNWALSVMYSYLEIYQ 189
QY 183 EKVLDLLDPAGDVLVIREDCRGNLILPGLSQKPISSFADFERHFLPASRNRVTGATRL 242
Db 190 EKVLDLLDPAGDVLVIREDCRGNLILPGLSQKPISSFADFERHFLPASRNRVTGATRL 249
QY 243 RSRSHAVLLVKVDQERLAPFQREGKLYLIDLAGSEDNRRTGNKGLRKESGAIN 302
Db 250 RSRSHAVLLVKVDQERLAPFQREGKLYLIDLAGSEDNRRTGNKGLRKESGAIN 309
QY 303 FVLGKVVADALNOGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDVT 362
Db 310 FVLGKVVADALNOGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDVT 369
QY 363 AARKEVINRPTNESLOPHALGPVKLSQKELLGPPAKARGPEEBEIGSPPEMAAPAS 422
Db 370 AARKEVINRPTNESLOPHALGPVKLSQKELLGPPAKARGPEEBEIGSPPEMAAPAS 429
```

QY 423 ASOKLSPLQKLSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRNRVLMKTVVEKDL 481
 Db 430 ASOKLSPLQKLSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRNRVLMKTVVEKDL 488

RESULT 4

US-09-925-300-1228
 ; Sequence 1228, Application US/09925300
 ; Patent No. US20020151681A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Craig Rosen,
 ; APPLICANT: Steve Ruben
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA101
 ; CURRENT APPLICATION NUMBER: US/09/925,300
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05988
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 1890
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1228
 ; LENGTH: 460
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (75)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (147)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (435)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; US-09-925-300-1228

Query Match 83.1%; Score 2151; DB 9; Length 460;
 Best Local Similarity 95.9%; Pred. No. 2.3e-172;
 Matches 424; Conservative 1; Mismatches 17; Indels 0; Gaps 0;
 QY 25 GAGCRLSKIGATRRPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANRNHQ 84
 Db 2 GAGCRLSKIGATRRPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANRNHQ 61
 QY 85 ETLYQPDFAFYGERSTOODIYAGSVQPIRLHLEGQNASVLYAGPTGAGKTHMLGSPEQ 144
 Db 62 ETLYQPDFAFYGERSTOODIYAGSVQPIRLHLEGQNASVLYAGPTGAGKTHMLGSPEQ 121
 QY 145 PGVPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRG 204
 Db 122 PGVPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRG 181
 QY 205 NILIPGLSQPISPFADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQRLAPF 264
 Db 182 NILIPGLSQPISPFADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQRLAPF 241
 QY 265 RQREGKLYLIDLAGSEDRRTGNKGLRLKESGAINSTLSFLVGLKVDALNOGLPRVYRDS 324
 Db 242 RQREGKLYLIDLAGSEDRRTGNKGLRLKESGAINSTLSFLVGLKVDALNOGLPRVYRDS 301
 QY 325 KLTRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPTNESLOPHAL 384
 Db 302 KLTRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPTNESLOPHAL 361
 QY 385 GPVKLSQKELLGPPPEAKRGPPEEEIGSPPEMAAPASQKLSPLQKLSMDPAMLERL 444
 Db 362 GPVKLSQKELLGPPPEAKRGPPEEEIGSPPEMAAPASQKLSPLQKLSMDPAMLERL 421
 QY 445 LSLDRLLASQSGQAPLLSTPK 466
 Db 422 LQLGPSACLPGEPPXPSVEYPK 443

RESULT 5

US-10-797-893-2
 ; Sequence 2, Application US/10797893
 ; Publication No. US20040142397A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; TITLE OF INVENTION: Novel motor proteins and methods for
 ; TITLE OF INVENTION: their use
 ; FILE REFERENCE: 1044
 ; CURRENT APPLICATION NUMBER: US/10/797,893
 ; CURRENT FILING DATE: 2004-03-09
 ; PRIOR APPLICATION NUMBER: US/09/724,224
 ; PRIOR FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: 09/597,292
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 370
 ; TYPE: PRT
 ; ORGANISM: Human
 ; US-10-797-893-2

Query Match 72.3%; Score 1873; DB 16; Length 370;
 Best Local Similarity 100.0%; Pred. No. 4.6e-149;
 Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 AAGSGTQORRREMAAASAAISGAGRCRLSKIGATRRPPPARVAVRLRPFDGTAGAS 62
 Db 2 AAGSGTQORRREMAAASAAISGAGRCRLSKIGATRRPPPARVAVRLRPFDGTAGAS 61
 QY 63 DPCVCGMDCSLEIANRNHQTLYQPDFAFYGERSTOODIYAGSVQPIRLHLEGQNA 122
 Db 62 DPCVCGMDCSLEIANRNHQTLYQPDFAFYGERSTOODIYAGSVQPIRLHLEGQNA 121
 QY 123 SVLAYGPTGAGKTHMLGSPEQGVIPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQ 182
 Db 122 SVLAYGPTGAGKTHMLGSPEQGVIPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQ 181
 QY 183 EKVLDDLPASGDLVIREDCRGNILIPGLSQKPISSPFADFERHFLPASNRRTVGATRLN 242
 Db 182 EKVLDDLPASGDLVIREDCRGNILIPGLSQKPISSPFADFERHFLPASNRRTVGATRLN 241
 QY 243 RSSRSHAVLLVKVDQRLAPFQREGKLYLIDLAGSEDRRTGNKGLRLKESGAINSTL 302
 Db 242 RSSRSHAVLLVKVDQRLAPFQREGKLYLIDLAGSEDRRTGNKGLRLKESGAINSTL 301
 QY 303 FVLGKVDALNOGLPRVYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNF 362
 Db 302 FVLGKVDALNOGLPRVYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNF 361
 QY 363 AARSKEVIN 371
 Db 362 AARSKEVIN 370

RESULT 6

US-10-797-893-6
 ; Sequence 6, Application US/10797893
 ; Publication No. US20040142397A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; TITLE OF INVENTION: Novel motor proteins and methods for
 ; TITLE OF INVENTION: their use
 ; FILE REFERENCE: 1044
 ; CURRENT APPLICATION NUMBER: US/10/797,893
 ; CURRENT FILING DATE: 2004-03-09
 ; PRIOR APPLICATION NUMBER: US/09/724,224
 ; PRIOR FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: 09/597,292
 ; PRIOR FILING DATE: 2000-06-20


```
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Human
US-10-797-893-6

Query Match      68.1%; Score 1763; DB 16; Length 346;
Best Local Similarity 100.0%; Pred. No. 7.8e-140;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 GCRLSKIGATRRPPARVAVRLRPFDVGTAGASDPPCVRGMDSCSLFIANNWNOET 86
DB 2 GCRLSKIGATRRPPARVAVRLRPFDVGTAGASDPPCVRGMDSCSLFIANNWNOET 61

QY 87 LKYPQDAFYGERSTQODIYAGSVQPIRLHLLBQNASVLAYGPTGAGKTHMLGSPQPG 146
DB 62 LKYPQDAFYGERSTQODIYAGSVQPIRLHLLBQNASVLAYGPTGAGKTHMLGSPQPG 121

QY 147 VIPRALMDLLQTRREGAEGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRNI 206
DB 122 VIPRALMDLLQTRREGAEGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRNI 181

QY 207 LIPGLSQKPISSFADFERHFLPASRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPRQ 266
DB 182 LIPGLSQKPISSFADFERHFLPASRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPRQ 241

QY 267 REGKLYLIDLAGSEDNRRGKGLRLKESGAINTSFLVLGKVVDALNOGLPRVPYRDSKL 326
DB 242 REGKLYLIDLAGSEDNRRGKGLRLKESGAINTSFLVLGKVVDALNOGLPRVPYRDSKL 301

QY 327 TRLQDSLGGSASHILIANIAPERRFYLDTSALNFAARSKVIN 371
DB 302 TRLQDSLGGSASHILIANIAPERRFYLDTSALNFAARSKVIN 346

RESULT 7
US-10-425-115-320209
; Sequence 320209, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 320209
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_55097C.1.pap
US-10-425-115-320209

Query Match      23.8%; Score 615.5; DB 17; Length 589;
Best Local Similarity 31.3%; Pred. No. 9e-43;
Matches 181; Conservative 95; Mismatches 169; Indels 133; Gaps 20;

QY 15 MAASAAATSGAGRCRLSKIGATRRPPARVAVRLRPFDVGTAGASDPPCVRGMDSC- 73
DB 1 MAASATPARSSA--HLSQ-----PVRVLRVRPFLSSEAAATAPCV-SLGGCH 46

QY 74 -----SLEIANWNRHOETKYQFDAPYGERSTQODIYAGSVQPIRLHLLBQNASVLAYG 128
DB 47 PGGGVTVQLKD-QHTSRSEQYKLDAPFQEDSVSQIFDQEVRAVIPSIFEGINATVFAYG 105

QY 129 PTGAGKTHMLGSPQPGVIPRALMDLLQTRREGAEGRPWALSVTMSYLEIYQEKVLDL 188
```

```
Db 106 ATGSKTYTMQGTEDFPGLIPLAASTILALCT-----GTWC-SVEISYEVYMERCDL 158
QY 189 LDPASGDLVIREDCRGNILIFGLSQKPISSFADFERHFLPASRRTVGATRLNQRSSRS 248
DB 159 LEFKAKEIMALDDDKGNQKGLSWPVRSMEEFQELYSIGVQRKKAHTGLNDVSSRS 218
QY 249 AVLKVDQERLAPRQERHFLPASRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPRQ 308
DB 219 AVLSLRVSTDVV-----KGKLNILIDLAGSEDNRRGKGLRLKESGAINTSFLVLGKV 272
QY 309 VDALNOGLPRVPYRDSKLTRLQDSLGGSASHILIANIAPERRFYLDTSALNFAARSK 368
DB 273 ISALNWNHRIPYRQSKLTRLRLDLSGGSRTVMIACLNPAB--YQESANTVSLAARSC 330
QY 369 VINRPTNESLQP-----HALGPVKLSQ----- 391
DB 331 IEN--FTSSSKQETPKLIDMEAKLRAWLESKGTKSQIRMDGLFSPIASKTPLSVSHMK 388
QY 392 -----KELLGP-----PEAKRARGPEEEIIGSPPEMAAP----- 420
DB 389 QPTSSRIPRCVKAMDQGGKIKILFDVHVHPTENI PREHROTQVNTPKVVLPSVTFC 448
QY 421 -----ASASOKLSPLOKLSMDPAMLERLLSLDRLASQSGCAPLL---STPKERMWL 472
DB 449 NEKHEASLRKALSPIS--SNMVP-----VEQOIPDNEN--CPVILLETPTIEKRNIV 496
QY 473 MKTVEEKDLE-IERLTKQKE-----LEAKMLAQKAE 503
DB 497 EETPGATPLERFNLGSLNKEALVQOYLDLLNVAKEE 534

RESULT 8
US-10-437-963-159031
; Sequence 159031, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 159031
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(377)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT4530_58448C.1.pap
US-10-437-963-159031

Query Match      23.2%; Score 601; DB 16; Length 377;
Best Local Similarity 36.7%; Pred. No. 7.7e-42;
Matches 151; Conservative 76; Mismatches 133; Indels 52; Gaps 12;

QY 31 LSKIGATRRPPARVAVRLRPFDVGTAGASDPPCVRGMDSCSLFIANNWNRHOETLK- 88
DB 1 MATAAATQSQP---RVVLRVRPPLPSAANGAEPVGLLGHGPGCVTVQLKQDYTSN 57
QY 89 --YQFDAPYGERSTQODIYAGSVQPIRLHLLBQNASVLAYGPTGAGKTHMLGSPQPG 146
```

Db 58 EYKLDAPFGQSRVCEIFDQEVSAVIGPIEGTNATVAYGATSGSKTYTMOCTEDLP 117
 Qy 147 VIPRALDMLLTREGEAGRPWALSVMYSLEYQEKVLDLLDPAGDLVIREDCRGN 206
 Db 118 LMXSTVALCTGT-----WC-SVEISYEVYMERCYDLEPKAREINVLDDKGNL 167
 Qy 207 LIPGLSOKPISPADFERHFLPASRNRVTGATRNQSRSSHAVLLVKVQORERLAPFR 266
 Db 168 QKGLAVPVRSLEPHEIYISIGVQRKVAHTGLNDVSSRSHAVLSIRITT-----DV 220
 Qy 267 REGKLYLI---DLAGSDNRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVYRD 323
 Db 221 VKGNLITXDLNAGSDNRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVYRD 280
 Qy 324 SKTRLLQDSIGGSAHSILIANIAPERRFVLDVTSALNFAARSKEVINRNPFTNBSLOPHA 383
 Db 281 SKLRLIQLDSIGGSHAVMLACLNPV---YQEAHTVSLAARSHTVN-----HM 328
 Qy 384 LQPVKLSQKLLGPPEAK-----RARGPEE-----EIGSGPEPMAAPASASQ 425
 Db 329 SSASKKXKDKVDM---EAKLRAWLESKGTKRSIQRMGDLSPNAIKTPLSNMQ 377

RESULT 9

US-10-437-963-103943
 ; Sequence 103943, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 103943
 ; LENGTH: 776
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_101324C.1.pep
 US-10-437-963-103943

Query Match 22.4%; Score 579; DB 16; Length 776;
 Best Local Similarity 30.6%; Pred. No. 1.6e-39;
 Matches 162; Conservative 100; Mismatches 209; Indels 58; Gaps 12;
 Qy 2 PAA-----GGSTQORREMAAASAAATSGAGRCRLSKIGATRPPPARVRVAVRLPFPVD 56
 Db 173 PAARKLSLGGGMAARLKTAGEAGNGDAAG-----SRINVFVRLRPMR 217
 Qy 57 GTAGADPPPCVRGWDSCSLBIANRHNQETLK-----YQDFAFYGERSTQODIYAGS 108
 Db 218 KEDAGSRSCVKIIVNKDVLYTFEASSETDYLRLKVRGRHFCFDSFPDPTTQAEVYST 277
 Qy 109 VQPIRLHLLGQNASVLAGPTCAGKTHMLGSPGQVPIPRALMDLLQITREGEAGRP 168
 Db 278 TSDLVGVLQGRNGTFCYCATGAGKTYTMLGTMSPGVMVLAIKDLFTKVRQRSHDGNH 337
 Qy 169 WALSVTMSYLEIYQEKVLDLLDPAGDLVIREDCRGNILPGLSOKPISGFADFERHFLP 228
 Db 338 ---SIQLSYLEVNEYTRDLSFGR-PLLRREDKQTV-AAGLTHYRAYSTDEVKLLQ 392
 Qy 229 ASNRVTGATRNQSRSSHAVLLVKVQORE-RLAPFRQREGKLYLIDLAGSDNRRTGN 287
 Db 393 GQNRRTTEPTFRVNETSSRSHAILQIVIVEYSIDGGSIVTVRGVGLSLIDLAGSERALATDQ 452

Qy 288 KGLRKESGAINTSLFVLGKVVDALNOGLPRVYRDSKLTLLQDSIGGSAHSILIANIA 347
 Db 453 RTQRESIGANINRSLALSCINALVEGKHIPRNSKLTQLLQDSIGGSCNTVMIANIS 512
 Qy 348 PERRFYLDVTSALNFAARSKEVINRNPFTNBSLOPHALGPVKLSQKELL-----GPPEAKR 402
 Db 513 PSNLSFGETQNTLHWADRAKEIKTKALTANEEVLRTVDSBTDQAKLVLELQKENSELRQ 572
 Qy 403 ARGPEEEISGPPEMAAPASASQKLSPLQKLSNDPAMLERLLSLDRILLASQSGGAPLL 462
 Db 573 QLARQQOKLLTVQAOTLASNASPOQSPASQAISTPCSTQKVRKRSTLAGN-----CF 625
 Qy 463 STPKREKRV-----LMKTVEKDELEIRLTKKQKELEAKMLAQKAE 503
 Db 626 NTFDSKRPAAENAQVRLDQKRVKAMEAEIEKMK-KEHLLQLK---QKDE 670

RESULT 10

US-10-425-115-332919
 ; Sequence 332919, Application US/10425115
 ; Publication No. US20040214272A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 332919
 ; LENGTH: 783
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(783)
 ; OTHER INFORMATION: unsure at all Xaa locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_66737C.1.pep
 US-10-425-115-332919

Query Match 22.3%; Score 578.5; DB 17; Length 783;
 Best Local Similarity 32.9%; Pred. No. 1.8e-39;
 Matches 167; Conservative 94; Mismatches 205; Indels 41; Gaps 13;
 Qy 11 RRREMAAASAAATSGAGRCRLSKIGA-TRPPPARVRVAVRLPFPVDGTAGSDPPPCVRG 69
 Db 180 RKLUSLGGVGGKUKAAG-----EVGAGTAAESRILVFLRPMGRKEKAGSRSCVKI 234
 Qy 70 MDSCSLBIANRHNQETLK-----YQDFAFYGERSTQODIYAGSVQPIRLHLLGQ 121
 Db 235 VNRKEVLTESASENDYLRLKGRDSDHFCFDSVPDSTQAEVYSTADLVEGVLOGRN 294
 Qy 122 ASVLAVGPTCAGKTHMLGSPGQVPIPRALMDLLQITREGEAGRPWALSVMYSYLEIY 181
 Db 295 GTVFCYCATGAGKTYTMLGTMTNPGVMVLAINDLFSKVTQKNH-----SIKLSYLEIY 347
 Qy 182 QEKVLDLLDPAGDLVIREDCRGNILPGLSOKPISGFADFERHFLPASRNRVTGATRN 241
 Db 348 NETVRDLSLPGS-PLNLRDQKQ-IVAAGLTQSRVSYSTDEVMLLOKGNKRTTEPTFRV 405
 Qy 242 QRSRSHAVLLVKVQORERLAPFRQREGKLYLIDLAGSDNRRTGNKGLRKESGAINTS 301
 Db 406 ETSRSHAVLVQVVEYRSLDGVNVKAGKLSLIDLAGSERALATDQRTQRSIEGANINRS 465
 Qy 302 LFLVGLKVVDALNOGLPRVYRDSKLTLLQDSIGGSAHSILIANIAPERFYLDVTSALN 361
 Db 466 LLALSSCINALVEGKHIPRNSKLTQLLQDSIGGACNTVMIANISPSNLSFGETQNTLH 525

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 Search time 10.0875 Seconds
(without alignments)
3300.235 Million cell updates/sec

Title: US-10-797-893-6
Perfect score: 1768
Sequence: 1 MGRCLSKIGATRRPPPARV.....FYLDTSALNFAARKEVIN 346

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1550	87.7	665	2 S62328	kinesin-like DNA b
2	594.5	33.6	664	2 T48258	kinesin-like prote
3	555.5	31.4	784	1 A55236	kinesin-related pr
4	550	31.1	703	2 A86319	F25116.11 protein
5	547	30.9	1254	2 T18277	kinesin heavy chai
6	536	30.3	784	2 T40594	probable kinesin-1
7	532	30.1	1226	2 I51617	kinesin-like prote
8	530.5	30.0	885	2 D86151	F22M8.8 protein -
9	527.5	29.8	1388	2 T30335	KLP2 protein - Afr
10	525	29.7	1070	2 T06733	kinesin homolog F2
11	524	29.6	1056	2 H84777	probable kinesin-r
12	523	29.6	1058	2 T47525	kinesin-related pr
13	521.5	29.5	1006	2 T02017	probable kinesin-1
14	515.5	29.2	1076	2 B84687	probable kinesin-1
15	514.5	29.1	699	1 S38982	kinesin-related pr
16	514	29.1	1066	1 A48669	kinesin-related pr
17	514	29.1	1695	2 A56921	kinesin family pro
18	513.5	29.0	701	1 B44259	kinesin-related pr
19	513.5	29.0	805	2 S64238	kinesin-related pr
20	513.5	29.0	1225	2 A56514	chromokinesin - ch
21	513	29.0	747	1 A57107	kinesin-related pr
22	511.5	28.9	1022	2 E84792	probable kinesin h
23	509.5	28.8	786	2 A53939	kinesin homolog KH
24	509.5	28.8	2954	2 T14156	kinesin-related pr
25	508	28.7	1229	2 T48959	kinesin-like prote
26	507.5	28.7	1231	2 A54803	microtubule-associ
27	507	28.7	928	2 T10164	kinesin heavy chai
28	504.5	28.5	330	2 B48935	kinesin-like prote
29	504	28.5	813	2 T46242	kinesin-like prote

30	501	28.3	1056	1 G02157	kinesin-like spind
31	500	28.3	1150	1 A55289	kinesin-like prote
32	498.5	28.2	935	2 T51930	kinesin (imported)
33	495.5	28.0	958	2 T20621	hypothetical prote
34	495	28.0	929	2 T51932	kinesin (imported)
35	491	27.8	963	1 A41919	kinesin heavy chai
36	490	27.7	742	1 S58691	kinesin-related pr
37	488	27.6	987	2 T51360	kinesin-like heavy
38	485.5	27.5	793	2 JC5831	kinesin-related pr
39	485.5	27.5	843	2 S44868	kinesin heavy chai
40	483	27.3	332	2 C48935	kinesin-like prote
41	482.5	27.3	883	2 T40128	kinesin-like prote
42	482.5	27.3	1584	1 JN0114	kinesin-related pr
43	482.5	27.3	1584	2 T15822	kinesin-like prote
44	482	27.3	1027	2 S37711	kinesin heavy chai
45	482	27.3	1031	1 A38713	kinesin heavy chai

ALIGNMENTS

RESULT 1

S62328
kinesin-like DNA binding protein KID - human
C:Species: Homo sapiens (man)
C:Date: 23-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 16-Aug-2004
C:Accession: S62328
R:Tokai, N.; Fujimoto-Nishiyama, A.; Toyoshima, Y.; Yonemura, S.; Teukita, S.; Inoue, J.; Embo, J. 15; 457-467, 1996
A:Title: Kid, a novel kinesin-like DNA binding protein, is localized to chromosomes and A:Reference number: S62328; MUID:96174806; PMID:8559929
A:Accession: S62328
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-665 <OK>
C:Superfamily: kinesin motor domain homology
C:Keywords: ATP; nucleotide binding; P-loop
F:44-374/Domain: kinesin motor domain homology <KMT>
F:128-135/Region: nucleotide-binding motif A (P-loop)

Query Match 87.7%; Score 1550; DB 2; Length 665;
Best Local Similarity 90.8%; Pred. No. 1.6e-117;
Matches 314; Conservative 1; Mismatches 29; Indels 2; Gaps 2;

Qy	2	GRCLSKIGATRRPPPARVAVRLRFVDTGATAGSDPPCVRGWMDSCSLEIANWRNHQET	61
Db	26	GRCLSKIGATRRPPPARVAVRLRFVDTGATAGSDPPCVRGWMDSCSLEIANWRNHQET	85
Qy	62	LKYQDFAFYGERSTQDDIYAGSVQPIRLHLEGGONAS-VLAYGPTGAGKTHMLGSPEQP	120
Db	86	LKYQDFAFYGERSTQDDIYAGSVQPIRLHLEGGONAKVILAYGPTGAGKT-THAQGP	144
Qy	121	GVTPRALMDLLQITREGAEGRPWALSVTVMSYLEIYQEKVLDLLDPASGDLVIREDCRGN	180
Db	145	WGDPAGSHGCPAAHKGCGCRGPAMGLSVTVMSYLEIYQEKVLDLLDPASGDLVIREDCRGN	204
Qy	181	ILIPGLSQKPISSFADFPERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPER	240
Db	205	ILIPGLSQKPISSFADFPERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPER	264
Qy	241	QREGKLYLDLAGSENNRRTGNKGLRKESGAINTSFLVLGKVVDALNQLGPRVPYRDSK	300
Db	265	QREGKLYLDLAGSENNRRTGNKGLRKESGAINTSFLGKVVVDALNQLGPRVPYRDSK	324
Qy	301	LTRLLOQSLGGSAAHSILIANIAPERFYLDTSALNFAARKEVIN	346
Db	325	LTRLLOQSLGGSAAHSILIANIAPERFYLDTSALNFAARKEVIN	370

RESULT 2

T48258
kinesin-like protein - Arabidopsis thaliana
N;Alternate names: protein T1E22.130

C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T48258
R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft, submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24489
A;Accession: T48258
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-664 <BEV>
A;Cross-references: UNIPROT:Q9LZ88; EMBL:AL162874
A;Experimental source: cultivar Columbia; BAC clone TIE22
C;Genetics:
A;Map position: 5
A;Introns: 58/3; 128/2; 184/3; 310/1; 378/3; 480/2; 501/3; 535/1; 607/3; 628/3
A;Note: TIE22_130

Query Match	33.6%;	Score	594.5;	DB	2;	Length	664;
Best Local Similarity	38.6%;	Pred. No.	4.2e-40;				
Matches	131;	Conservative	69;	Mismatches	114;	Indels	25;
				Gaps	6;		
Qy	18	ARYVAVRLRPV	-----DGTAGADPPCVRGWDSCSLEIANWRNHQETLK----	YQFDA	68		
Db		: : : : : : : : :	: : : : : : : : :				
Qy	78	SNRVVLRVLRPF	PREISDESCDGRSCVSVIGDGGDTSEVAVYLKDPDSCRNESYQLDA	137			
Db		: : : : : : : : :	: : : : : : : : :				
Qy	69	FGY-ERSTQDIYAGS	QPTLRHLLEGQNASVLAYGPTGAGKTHTMLGSPQGVPIPRAL	127			
Db		: : : : : : : : :	: : : : : : : : :				
Qy	138	FGYGEDNVKHI	FDREVSPILPGIFHGFNATVLAYGATGSGKFTWQGIDELPLMLPTM	197			
Db		: : : : : : : : :	: : : : : : : : :				
Qy	128	MDLLQITREIGASGR	PWALSVMTVSYLEIYQEKVLDDLDPASGDLVIREDCRGNILIPGLS	187			
Db		: : : : : : : : :	: : : : : : : : :				
Qy	198	STILSMCEKTRSAE	-----ISYVEVYMDRCWDLLEVKDNEIAVDDKDGQVHLKGLS	250			
Db		: : : : : : : : :	: : : : : : : : :				
Qy	188	QKSISSFADFERHFL	PASRNRTYGTALNQBSRSHAVLLVKVDQERLAPFRQREGKLY	247			
Db		: : : : : : : : :	: : : : : : : : :				
Qy	251	SVPVKMGSEFOEAYL	CGVQRKKVAHTGLNDVSSRSHGVLVISVTSQGLV-----GKIN	304			
Db		: : : : : : : : :	: : : : : : : : :				
Qy	248	LIDLAGSEDNRRTGN	KGLRLKESGANTSLFVLGKVVDALNQLPVPVPYRDSKLTLLQD	307			
Db		: : : : : : : : :	: : : : : : : : :				
Qy	305	LIDLAGNEDNRTEG	IRLOESAKINQSLFALSNNVYALNNNLPRVPVRETAKLTILQD	364			
Db		: : : : : : : : :	: : : : : : : : :				
Qy	308	SLGCSAHSILIANIA	PERRFYLDTVSALNPAARKEVIN	346			
Db		: : : : : : : : :	: : : : : : : : :				
Qy	365	SLGTSRALVACIN	PGCE--YOESLRTVSLAASRSHIN	401			
Db		: : : : : : : : :	: : : : : : : : :				

RESULT 3
A55236
N;Alternate names: kinesin-like protein 5; KLP5
C;Species: Drosophila melanogaster
C;Date: 23-Mar-1995 #sequence.revision 23-Mar-1995 #text_change 09-Jul-2004
R;Accession: A55236; E41298
R;Pesaavento, P.A.; Stewart, R.J.; Goldstein, L.S.B.
J. Cell Biol. 127, 1041-1048, 1994
A;Title: Characterization of the KLP68D kinesin-like protein in Drosophila: possible rol
A;Reference number: A55236; MUID:95050960; PMID:7525600
A;Accession: A55236
A;Molecule type: mRNA
A;Residues: 1-784 <PES>
A;Cross-references: UNIPROT:P46867; GB:U15974; NID:G595912; PIDN:AAA69929.1; PID:G565090
R;Stewart, R.J.; Pesaavento, P.A.; Woerpel, D.N.; Goldstein, L.S.B.
Proc. Natl. Acad. Sci. U.S.A. 88, 8470-8474, 1991
A;Title: Identification and partial characterization of six members of the kinesin super
A;Reference number: A41298; MUID:92020874; PMID:1924306
A;Accession: E41298
A;Molecule type: DNA
A;Residues: 'TC', 222-337, 'VRGVQV' <STE>
A;Cross-references: GB:W7431; NID:G157791; PIDN:AAA28658.1; PID:G157792
C;Genetics:
A;Gene: FlyBase:Klp68D; KLP5
A;Cross-references: FlyBase:FBgn0004381
C;Function:

A,Description: may be part of a motor protein that provides anterograde fast axonal transport.
C:Superfamily: kinesin-related protein KIF; kinesin motor domain homology
C:Keywords: ATP; coiled coil; microtubule binding; nucleotide binding; P-Loop
F:1-349/Domain: head globular #status predicted <GL>
F:20-350/Domain: kinesin motor domain homology <KMT>
F:106-113/Region: nucleotide-binding motif A (P-loop)
F:350-580/Domain: helical rod #status predicted <RD>
F:581-784/Domain: tail globular #status predicted <TGL>
F:112/Binding site: ATP (Lys) #status predicted

Query Match	31.4%;	Score 555.5;	DB 1;	Length 784;
Best Local Similarity	39.5%;	Pred. No. 7.7e-37;		
Matches 140;	Conservative 56;	Mismatches 127;	Indels 31;	Gaps 10;

Qy	12	TRRP-----PPARVRVAVLRPPFVDGTAGADPPCV-----RGMDSCSLEIANW--RN	57
Db	5	SRRPFGSSGOTPNECVQVVVRCPMSNRERSRSEPVVNVYPNRGV----VELQNVVDGN	60
Qy	58	HQBTLKKQFPAFYGERSTQODIYAGSVQPTLRHLLEGQNASVLAYGPTGAGKHTM--L	114
Db	61	KEQKVFTYDAADASATQTTLHEVVFPVLVSSVLEGFNGCIFYAGTGTTGKTFMGVR	120
Qy	115	GSPEQPGVPIPRALMDL-LQITREBEGAGRPWALSVTMSYLEIYQEKVLDLLDPASGD	173
Db	121	GNDELGMIIPTTEQIWLHINRTEN-----FQFLADVSVYLEIYMEELRDLKKPNSKH	175
Qy	174	REDCRGN-ILI PGLSQPKISSPADFERHFHPASRNRTVGATRLNQRSSRSHAVLLVK	232
Db	176	RE--RGSVVVPVNLHAINCKSVEDMTKVMQVGNKRTVGFNTNNNEHSRSHAFPMIK	233
Qy	233	RERLAPRQREGKLYLIDLAGSDNRRTNGKLRLKESGAINTSFLVLGKVVADALNOGLP	292
Db	234	CD--TETNTYIKVGNLIDLAGSRQSKTGASARELKEASKINLALSISLGNVISALAES	292
Qy	293	RVPYRDSKLTRLLODSLGGSAHSILIANIAPERRFYLDTVTSALNFAARSKEVIN	346
Db	293	HVPYRDSKLTRLLODSLGGSKTMIANIGFSNNVNYNETLTLTYRYSGRASIKON	346

RESULT 4
A86319
P25116.11 protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: A86319
R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; D.
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.;
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.;
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A86319
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-703 <STO>
A:Cross-references: UNIPROT:Q9FZ77; GB:AE005172; NID:g9795601; PIDN:AAF98419.1;
C:Genetics:
A:Map position: 1

	Query Match	31.1%; Score 550; DB 2;	Length 703;
	Best Local Similarity	38.6%;	Pred. No. 1.8e-36;
	Matches 130; Conservative	62; Mismatches 113;	Indels 32; Gaps 6;
Qy	18 ARRVAVARLRPPFDVGTAGASDPFCVRGMDSCSLEIANRNRHQETLK-----YQDAF	69	
	: :: :	:::~::~:	:
Dd	150 SLTVLFRVLRLPMGGKERENGSRCCVKAKRDLYLTFTNENDYLRLKRLVRHRTDSS	209	
	: :: :	:::~::~:	:
Qy	70 YGRSTCOODIYAQSVOPIILNLHLLEGQWASVLAYGPTCAGKTHTMWLGSPGPVPTRALMD	129	

Db 210 FPETTOQEVSTTGDVLEAVLEGRNSVFCYGATGAGKTYTMTGMENFGVMVLAIKD 269
Qy 130 LLQLTRBEGAGRPWALSVTMSYLEIYOEKVLDDLDPASGLVIREDCRGNILIPGLSQK 189
Db 270 LFAKVRQSLDGNH---VVHLSYLEVNETVRDLSLSPOR-PLIUREDKQ-----VWALLQR 321
Qy 190 PISSFADPERHFLPASNRRTVGATRLNQSRSSHAVLLVKVDORERLAPFR--OREGKLY 247
Db 322 -----GNQNRTEPTRCNETSSRSRSHAILQVIVEYKTRDASNNIISRVGKLS 367
Qy 248 LIDLAGEDNRRTGNKGLRKESGAINTSFLVLGKVDALNOGLPRVPYRDSKLTLLQD 307
Db 368 LIDLAGEDNRRTGNKGLRKESGAINTSFLVLGKVDALNOGLPRVPYRDSKLTLLQD 307
Qy 308 SLGSGSAHSILIANIAPERRFLDTVTSALNPAARKEV 344
Db 428 SLGSGCNTVMIANISPSQSFGETQNTLHWADRAKEI 464

RESULT 5

T18277
kinesin heavy chain - slime mold (Dictyostelium discoideum)
C/Species: Dictyostelium discoideum
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T18277
R/DeHostos, E.L.; McCaffrey, G.; Vale, R.D.
submitted to the EMBL Data Library, November 1995
A/Reference number: Z18953
A/Accession: T18277
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1254 <DE>
A/Cross-references: UNIPROT:Q94463; EMBL:U41289; NID:g1526990; PID:g1526991; PIDN:AA077

Query Match 30.9%; Score 547; DB 2; Length 1254;
Best Local Similarity 36.3%; Pred. No. 7.2e-36;
Matches 122; Conservative 76; Mismatches 126; Indels 12; Gaps 5;

Qy 15 PPARVAVRLRPVFDGTAGASPPCVRGWDSLSRI-ANWRNHQETLKYQPDAGFR 73
Db 24 PVSNIIVCVRLPTELEGRNHSIVHFDKSKSIRANGP-----QTFDRIFGY 77
Qy 74 STQDIYAGSVQPIRLHLEQNASVLAYGPTGAGKTHMLGSPQGVIPRALMDL--- 130
Db 78 ETQSQIPEDVAEPIVNDFLDGYHCTIAYGQTASGKFTWVGDPDSHGIIIPRVIESIFVG 137
Qy 131 LQLTRBEGAGRPWALSVTMSYLEIYOEKVLDDLDPASGLVIREDCRGNILIPGLSQK 190
Db 138 ISKMRKEDT-SLSLAFCLKISALELYNEKLYDLYIASKSNLNTREHKQNGIYVEGISEIV 196
Qy 191 ISSFADPERHFLPASNRRTVGATRLNQSRSSHAVLLVKVDORERLAPFRQREGKLYLD 250
Db 197 ITSIEEYFNLTISNNRRASTQMSAASRSHSLVMIELSQ-QNLSMWSKSKULFLVD 255
Qy 251 LAGSEDRNRRTGNKGLRKESGAINTSFLVLGKVDALNOGLPRVPYRDSKLTLLQDLSLG 310
Db 256 LAGSDISHKTGAEGDRQAEAKNINLSALGKVINALTGCVNYVPYRDSKLTTRVLQDLSLG 315
Qy 311 GSAHSILIANIAPERRFLDTVTSALNPAARKEVIN 346
Db 316 GNSKTSLIINCSPSNNNEHETITTLQFGTRAKTIKN 351

RESULT 6

T40594
probable kinesin-like protein - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T40594; T39531
R/Purnelle, B.; Goffeau, A.; Lyne, M.; Wood, V.; Rajandream, M.A.; Barrrell, B.G.
submitted to the EMBL Data Library, May 1998
A/Reference number: Z21939

A/Accession: T40594
A/Molecule type: DNA
A/Residues: 1-784 <PUR>
A/Cross-references: UNIPROT:O59751; EMBL:AL023587; PIDN:CAA19043.1; GSPDB:GN00067; SPDB:SPBC1685.15C
R/Wood, V.; Rajandream, M.A.; Barrrell, B.G.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, July 1998
A/Reference number: Z21861
A/Accession: T39531
A/Molecule type: DNA
A/Residues: 464-784 <WOO>
A/Cross-references: EMBL:AL031154; PIDN:CAA20063.1; GSPDB:GN00067; SPDB:SPBC1685.15C
A/Experimental source: strain 972h-; cosmid c1685
C/Genetics:
A/Gene: SPDB:SPBC649.01c; SPDB:SPBC1685.15C
A/Map position: 2
A/Introns: 11/1

Query Match 30.3%; Score 536; DB 2; Length 784;
Best Local Similarity 35.5%; Pred. No. 2.9e-35;
Matches 139; Conservative 61; Mismatches 116; Indels 76; Gaps 12;

Qy 20 VRVAVRLRPVFDGTAG--ASDPCCVRGMDSCSLSIANRNH----- 58
Db 7 ISVAVRVRPPTEREKGLLAETPKSKEFLGDSLAVSNTSNTCTGTNGIRKIVRVLDNDVL 66
Qy 59 -----QETL-----KYQFDAPYGERSTOODIYAGSVQPIRLHLEGG 95
Db 67 IFDPPENPLAKVQKSLLPAGKPRDVRVYAFDRULFGSEASQSDVYKGTTEPLDSVLQGY 126
Qy 96 NASVLAYGPTGAGKTHMLGSPQGVIP---PRALMDLQLTRBEGAGRPWALSVTMSY 152
Db 127 NATVFAVGATCGKTHITISRPDPDGIIFLTMRALD-----RVGLK-RTMNVDIVSVY 180
Qy 153 LEIYQEKVLDL---DPASGD-----LVIREDCRGNILIPGLSQKPISSFADPERHFLPASR 206
Db 181 LEIYNEKIRDLVLQDPLSMKPKSLNICEAEQNVSVFGLSYFTPTNLEEVMEIIRGNS 240
Qy 207 NRTVGATRLNQSRSSHAVLLVKVDORERLAPFRQREGK-----LVYLIDLAGSD 256
Db 241 NRTWSPTEANAVSRSHAVLQIYITQPKSGE-KQEESESONSHKVSFSDLAGSER 299
Qy 257 NRTTGNKGLRKESGAINTSFLVLGKVDALNOGLPR-----VPYRDSKLTLLQDLSLGS 312
Db 300 ASATKVRGKELVEGANINRSLLALNGCINSLSCE--PRRQHVVPYRDSKLTLLKFSLLGN 357
Qy 313 AHSILIANIAPERRFLDTVTSALNPAARKEV 344
Db 358 CRTCMIVCISPSSEHYDETHNTLKYGNRAKNI 389

RESULT 7

I51617
kinesin-like protein 1 - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Aug-2004
C/Accession: I51617; A48835; S48837
R/Vernos, I.; Raats, J.; Hirano, T.; Heasman, J.; Karsenti, E.; Wylie, C.
Cell 81, 117-127, 1995
A/Title: Xklp1, a chromosomal Xenopus kinesin-like protein essential for spindle organi
A/Reference number: A56221; MUID:95236444; PMID:7720067
A/Accession: I51617
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1226 <VER>
A/Cross-references: UNIPROT:Q91784; EMBL:X82012; NID:g562792; PIDN:CAA57539.1; PID:g562
R/Vernos, I.; Heasman, J.; Wylie, C.
Dev. Biol. 157, 232-239, 1993
A/Title: Multiple kinesin-like transcripts in Xenopus oocytes.
A/Reference number: A48835; MUID:93246085; PMID:8482413
A/Accession: A48835
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA

T06733
kinesin homolog F28P10.150 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T06733
R:Quetier, F.; Choine, N.; Robert, C.; Brottier, P.; Wincker, P.; Cantolico, L.; Artigou
submitted to the Protein Sequence Database, April 1999
A:Reference number: 215793
A:Accession: T06733
A:Molecule type: DNA
A:Residues: 1-1070 <QUE>
A:Cross-references: UNIPROT:Q9SV36; EMBL:AL049655; GSPDB:GN00061; ATSP:F28P10.150
A:Experimental source: cultivar Columbia; BAC clone F28P10
C:Genetics:
A:Gene: ATSP:F28P10.150
A:Map position: 3
A:Introns: 113/3; 256/3; 330/3; 377/1; 427/1; 456/3; 527/3; 572/3; 632/3;
C:Superfamily: kinesin-related protein KLP61F; kinesin motor domain homology
F;124-466/Domain: kinesin motor domain homology <KNOT>
Query Match 29.7%; Score 525; DB 2; Length 1070;
Best Local Similarity 35.3%; Pred. No. 3.5e-34;
Matches 126; Conservative 69; Mismatches 120; Indels 42; Gaps 10;
QY 17 PARVRVAVLRPFVDGTAGASD-----PPCVRGMDSCSLEIANWRNHQETLKYQFD 67
DB 121 PGRVRSVRVRP-RNGEELISDADFADLVELQPEIKRL-----KLKKNWNS-----SYKFD 172
QY 68 AFGERSTQDIYAGSVOPILRHLLLEGONASVLAAYPTGAGKTHM--LGSPR--OPGVI 123
DB 173 EVFTDTSSQKRVYEGVAKPVVEGLSYNGTINMAYGQTGTGKTYTGKIGKDDAAERGIM 232
QY 124 PRALMDLLQLTREEGAGRPWALSVMYSLEIYQEKVLDLDPASGDLVIREDCR--GNIL 182
DB 233 VRALEIL-----LNASSASISVEISYLQVETIQLLAPEKKNISINEDAKTGEVS 285
QY 183 IGLSKQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRAVLLVKV-----DQERILA 237
DB 286 VPGATVNIQDLDFLQVLQVGTNRHAANKNTNTESSRAHLLTVVRRAMNEKTEKAK 345
QY 238 PF-----RQREGKLYLIDLAGSDNRRTGKGLRLKESGAINTSLFVLGKVVDALNQ 289
DB 346 PESLGDKAIPRVKSKLLIYDLGSRINKSGTGDHMEAEKFINSLTSLGKCNALAE 405
QY 290 GLPRVPRYDSKLTLLQDLSGSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 346
DB 406 GSSHIPTRUSKLTLLRDSFGGARSARTSLIITIGPSARYHAETTTIMFGORAMKIVN 462
RESULT 11
H84777
probable kinesin-related cytochrome protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: H84777
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: H84777
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-1056 <STO>
A:Cross-references: UNIPROT:P82266; GB:AE002093; NID:g4510356; PIDN:AAD21445.1; GSPDB:GN
C:Genetics:
A:Gene: At2g36200
A:Map position: 2
C:Superfamily: kinesin-related protein Eg5; kinesin motor domain homology
Query Match 29.6%; Score 524; DB 2; Length 1056;
Best Local Similarity 36.4%; Pred. No. 4.1e-34;

Matches 131; Conservative 67; Mismatches 118; Indels 44; Gaps 9;
QY 20 VRVAVLRPFVDGTAGASDPPCVRGMDSCSLEIANWRN---HOETLKYQFDAPYGRST 75
DB 13 VQVLLACRFSDDEL-RSNAPQVLTCDLQREVAVSQNTAGKHIDRV-FTFDKVFGPSAQ 70
QY 76 QDIYAGSVOPILRHLLLEGONASVLAAYPTGAGKTHMLGS-----PEQPGVI 123
DB 71 QKDLVQAVVPIVNEVLEGFNCTIFAYGQTGTGKTYTMEGECRRSKSAPCGGLPAEAGVI 130
QY 124 PRALMDLLQLTREEGAGRPWALSVMYSLEIYQEKVLDLDP-----ASGDIV 172
DB 131 PRAVKQIF-----DTLEGQQAQYVSKVTFLELYNEEITDLLAPEDLSRVAERKQKKPLP 185
QY 173 IREDCRGNILIPGLSKQPISS---PADFERHFLPASNRRTVGATRLNQRSSRSRAVLLV 228
DB 186 LMEDGKGVLRGLEEIVTSANEIFTLER-----GSSKRTAETFLNKQSSRSLSFSI 241
QY 229 KVDQERERLAPFRQ--REGKLYLIDLAGSDNRRTGKGLRLKESGAINTSLFVLGKVYDA 286
DB 242 TTHKEATPEGEELIKCGKLNLDLAGSENISRGSGARDGRAREAGINKSLTLGRVISA 301
QY 287 LNQGLPRVPRYDSKLTLLQDLSGSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 346
DB 302 LVEHLGHVPRYDSKLTLLRDSLGSRKTKTCTIATVSPAVHCEETLSTLDYAHRAKNIRN 361
RESULT 12
T47525
kinesin-related protein-like - Arabidopsis thaliana
N:Alternate names: protein F16L2.60
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C:Accession: T47525
R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Rudd, S.;
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24468
A:Accession: T47525
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1058 <JOR>
A:Cross-references: UNIPROT:O9LZU5; EMBL:AL162459
A:Experimental source: cultivar Columbia; BAC clone F16L2
C:Genetics:
A:Map position: 3
A:Introns: 56/2; 99/3; 151/3; 310/3; 343/2; 395/3; 419/1; 444/3; 462/3; 489/3; 527/1; 5
A:Note: F16L2.60
C:Superfamily: kinesin-related protein KIP1; kinesin motor domain homology
Query Match 29.6%; Score 523; DB 2; Length 1058;
Best Local Similarity 33.4%; Pred. No. 5e-34;
Matches 126; Conservative 74; Mismatches 127; Indels 50; Gaps 8;
QY 5 RLKSGKATRRPPPAR---VRVAVLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61
DB 31 RSSESSTNRNDKEGVNVQVILRCPLSEDEARHTPVVI---SCN-----ENRREV 80
QY 62 LKQY-----FDYFGERSTQDIYAGSVOPILRHLLLEGONASVLAAYPTGAGK 109
DB 81 AATQSIAGKHIDRHFAFDVFGPASQKQDLYQAIQVIFVEVLEGYNCTIFAYGQTGTGK 140
QY 110 THTMLGS-----PEQPGVPIPRALMDLLQLTREEGAGRPWALSVMYSLEIYQEKVL 161
DB 141 TYTMEGARKNGEPPSDAGVPRAVKQIFDLEAOGAE-----YMKVTFLELYNEEIS 195
QY 162 DLLDP-----ASGDLVIREDCRGNILIPGLSKQPISSPADFERHFLPASNRRTVG 211
DB 196 DLLAPEETIKFVDEKSKKIALMEDGKGSFVRGLEEEIVSTANEIYKILEKSAKRRTA 255
QY 212 ATRLNQRSSRAVLLVKVQDQERERLAPFRQ--REGKLYLIDLAGSDNRRTGKGLRLKE 269
DB 256 ETLNKKQSSRSISFTIHKENTPEGEEMIKCGKLNLDLAGSENISRGSGAREGRARE 315

QY 270 SGAINSTSLFVLGKVVADLNOGLPRVPRDLSKTLRLQDSLGSAHSILIANIAPERRFVL 329
 Db 316 AGEINKSLTLGRVINALVEHSHIPYRDSKTLRLSLGKTKTCVIATISPSIHCL 375
 QY 330 DTVSALNFAARKEVIN 346
 Db 376 ETLSTLDYAHRAKNKN 392

RESULT 13
 T02017
 Kinesin-related protein TKRP125 - common tobacco
 C:Species: Nicotiana tabacum (common tobacco)
 C>Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 09-Jul-2004
 C:Accession: T02017
 R:Asada, T.; Kuriyama, R.; Shibaoka, H.
 J. Cell Sci. 110, 179-189, 1997
 A:Title: TKRP125, a kinesin-related protein involved in the centrosome-independent organ
 A:Reference number: Z14490; MUID:97196959; PMID:9044048
 A:Accession: T02017
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1006 <ASA>
 A:CROSS-references: UNIPROT:023826; EMBL:D83711; NID:g2582970; PIDN:BAA23159.1; PID:g258
 C:Genetics:
 A:Note: TKRP125
 C:Superfamily: kinesin-related protein Egs; kinesin motor domain homology
 F:10-361/Domain: kinesin motor domain homology <KMT>

Query Match 29.5%; Score 521.5; DB 2; Length 1006;
 Best Local Similarity 35.6%; Pred. No. 6.2e-34;
 Matches 128; Conservative 69; Mismatches 118; Indels 45; Gaps 9;

QY 20 VRVAVLRPFVDTAGASDPPCVRGMD-----SCSLEIANWRNHQETLKYQDFAFYGRS 74
 Db 10 VQVLLRCRPFSDNLRNAPVVTGNDYQREVAVSQNIAG--KHIDRI-FTEDKVGPSA 66
 QY 75 TQDIYAGSVQPIRLHLLGQNASVLAYGPTGAGTKHTMLGS-----PEQGV 122
 Db 67 QQRDLVDQAIIVNEVLEGFNCTIPAYGQTGTGTYMEGCKRSKGPNGELPQEA 126
 QY 123 IPRALMDLQLTREGAEGRPWALSVTMSYLEIYQEKVLDLDP-----ASGDIV 172
 Db 127 IPRAVKQVDTLESQAE-----YSVKVTFLELYNEETDILLAPEDLVKVALEDROKKQLP 181
 QY 173 IREDCRGNLIPGLSQKPTSS----PADFERHFLPASNRRTVGATRLNQRSSRSHAVLV 228
 Db 182 LMEDGKGVLRGLBEEIVTSANEIPTLLER----GSAKRRRTAETLLNQRSSRSHLSFI 237
 QY 229 KYDQERLAPFRQ--REGKLYLIDLAGSEDNRRTGNKGLRKESGAINSTSLFVLGKVDA 286
 Db 238 TTHIKETAPGEBELIKCGKLNLDVLAGSENISRSAGREAREAGEINKSLTLGRVINA 297
 QY 287 LNQGLPRVPRDLSKTLRLQDSLGSAHSILIANIAPERRFVLDTVSALNFAARKEVIN 346
 Db 298 LVEHLGHIPYRDSKTLRLDLSLGGTKTCIIATVSPAVHCLTEETSLDLYAHRAKNKN 357

RESULT 14
 B04687
 probable kinesin-like spindle protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
 C:Accession: B04687
 R:Lin, X.; Kaul, S.; Rounale, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: B04420; MUID:20083487; PMID:10617197
 A:Accession: B04687
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-1076 <STO>
 A:CROSS-references: UNIPROT:Q9STB3; GB:AE002093; NID:g4580395; PIDN:AAD24373.1; GSPDB:G
 C:Genetics:
 A:Gene: At2G28620
 A:Map position: 2
 C:Superfamily: kinesin-related protein KIP1; kinesin motor domain homology

Query Match 29.2%; Score 515.5; DB 2; Length 1076;
 Best Local Similarity 35.0%; Pred. No. 2.1e-33;
 Matches 122; Conservative 68; Mismatches 132; Indels 27; Gaps 6;

QY 20 VRVAVLRPFVDTAGASDPPCVRGMDSCSLEIA---NWRNHQETLKYQDFAFYGERSTQ 76
 Db 51 IQVIVRCRPF--NSETRIQTFAVLTCNRKKEVAQAQNIAGKIDKTLFLDKVFGPTSOQ 109
 QY 77 QDIYAGSVQPIRLHLLGQNASVLAYGPTGAGTKHTMLGS-----PEQGVIPRALM 128
 Db 110 KDLYHQAVSPVFEVLGYNCTIPAYGQTGTGTYMEGARKKNGEIPSDAGVIPRAVK 169
 QY 129 DLLQLTREGAEGRPWALSVTMSYLEIYQEKVLDL-----DPASGDIVIREDCRG 179
 Db 170 QIFDILEAQSA---EYSLKVSFLELYNEELTDLLAPEETKFADDKSKKPLALMEDGKG 225
 QY 180 NILIPGLSQKPTSSPADFERHFLPASNRRTVGATRLNQRSSRSHAVLVKVDORERLAPF 239
 Db 226 GVFRGLBEEIVSTADEIYKVKESAKRRTAETLLNQRSSRSHSIFSVTHIKECTPEG 285
 QY 240 RQ--REGKLYLIDLAGSEDNRRTGNKGLRKESGAINSTSLFVLGKVVDALNQGRLPRVYR 297
 Db 286 EEIVKSGKLNLDVLAGSENISRSAGREAREAGEINKSLTLGRVINALVEHSHIPYR 345
 QY 298 DSKLTRLQDSLGSAHSILIANIAPERRFVLDTVSALNFAARKEVIN 346
 Db 346 ESKLTRLRLDLSGKTKTCVIATVSPVHCLTEETSLDLYAHRAKHKN 394

RESULT 15
 S38982
 Kinesin-related protein KRP85 - sea urchin (Strongylocentrotus purpuratus)
 N:Alternate names: kinesin-2 chain A; KRP (85/95) 85K chain
 C:Species: Strongylocentrotus purpuratus (purple urchin)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C:Accession: S38982; S72551
 R:Cole, D.G.; Chinn, S.W.; Wedaman, K.P.; Hall, K.; Vuong, T.; Scholey, J.M.
 Nature 366, 268-270, 1993
 A:Title: Novel heterotrimeric kinesin-related protein purified from sea urchin eggs.
 A:Reference number: S38982; MUID:94050179; PMID:8232586
 A:Accession: S38982
 A:Molecule type: mRNA
 A:Residues: 1-699 <COL1>
 A:CROSS-references: UNIPROT:P46872; EMBL:L16993; NID:g295245; PIDN:AAA16098.1; PID:g2952
 A:Accession: S72551
 A:Molecule type: protein
 A:Residues: 2-5,'X',7-11,59-64,125-132,222-226,'X',228-230 <COL2>
 C:Complex: heterotrimer of a 115K chain and two kinesin-related chains of 95K (PIR:S5869
 C:Superfamily: kinesin-related protein KIF3; kinesin motor domain homology
 C:Keywords: ATP; heterotrimer; microtubule binding; nucleotide binding; P-loop
 F:11-348/Domain: kinesin motor domain homology <KMT>
 F:97-104/Region: nucleotide-binding motif A (P-loop)
 F:103/Binding site: ATP (lys) #status predicted

Query Match 29.1%; Score 514.5; DB 1; Length 699;
 Best Local Similarity 37.6%; Pred. No. 1.4e-33;
 Matches 127; Conservative 61; Mismatches 135; Indels 15; Gaps 8;

QY 20 VRVAVLRPFVDTAGASDPPCVRGMDSC--SLEIANWR--NHQETLKYQDFAFYGERST 75
 Db 11 VRVVRCPNLNKEGTGQGFKSVK--MDEMRTGVQVNTNAPSGEPKSFPTDTPAPGAK 69
 QY 76 QDIYAGSVQPIRLHLLGQNASVLAYGPTGAGTKHTMLG---SPEQGVIPRALMDLQ 132
 Db 70 QTDVYNQATPIVDAILIEGYNGTIPAYGQTGTGTYMEGVRSQPELRLIIPNSFAHIFG 129

```

Qy 133 LTREBAGRPMALSVTMSYLEIYOEKVLDLL-DPASGDLVIREDCRGNILIPGLSQKPI 191
Db 130 HIAKEQENVR---FLVRVSYLEIYNEEVKDLGKQQHRLQVKEKRPDGVVVKDLSAFV 186
Qy 192 SSFADFERHFLPASRNETVGATRLNQRSSRSHAVLLVKVDQERLAPFRQ--REGKLYLI 249
Db 187 NNADDMDRIMTLGNKNRSVGATNNWSSSRSHAFTITLERSDMGLDKEQHVRVYGLHMV 246
Qy 250 DLAGEDNRRTGNKGLRLKESGAINTSLFVLGKVVDALNOQ-LPRVPYRDSKLTLLQDS 308
Db 247 DLASERQTKTGATGQRLKEATKINLSLSTLGNVISSLVGDKSTHIYRNSKLTLLQDS 306
Qy 309 LGSASHILIANIAPERFYLDTVSALNFAARSKEVIN 346
Db 307 LGGNAKTMVCANIGFAEYNYDETISTLRYANRAKNIKN 344

```

Search completed: November 5, 2004, 18:45:33
Job time : 11.0875 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 ; Search time 71.0157 Seconds
(without alignments)
1747.786 Million cell updates/sec

Title: US-10-797-893-6

Perfect score: 1768

Sequence: 1 MGRCLSKIGATRRPPPARV.....FYLDTSALNFAARSEKVIN 346

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1768	100.0	346	7 ADC23342	Adc23342 Human kin
2	1768	100.0	346	8 ADQ60232	Adq60232 Human mic
3	1768	100.0	487	7 ADC23344	Adc23344 Human kin
4	1768	100.0	487	8 ADQ60234	Adq60234 Human mic
5	1763	99.7	370	7 ADC23338	Adc23338 Human kin
6	1763	99.7	370	8 ADQ60228	Adq60228 Human mic
7	1763	99.7	490	7 ADC40973	Adk40973 Novel hum
8	1763	99.7	512	7 ADC23340	Adc23340 Human kin
9	1763	99.7	512	8 ADQ60230	Adq60230 Human mic
10	1763	99.7	665	8 ADQ09240	Adq09240 Human kin
11	1753	99.2	460	3 AAB56650	Aab56650 Human pro
12	559.5	31.6	784	4 ABB71112	Abb71112 Drosophil
13	559.5	31.6	548	7 ADm04007	Adm04007 Human pro
14	559	31.6	998	8 AD044167	Ado44167 Structura
15	552.5	31.2	357	5 AAU76958	Aau76958 Novel hum
16	552.5	31.2	898	4 ABUS3123	Abu3123 Intracell
17	552.5	31.2	898	5 AAU76957	Aau76957 Novel hum
18	552.5	31.2	898	7 ADC31082	Adc31082 Human nov
19	552.5	31.2	898	7 ADI15915	Adi15915 Human pp
20	551.5	31.2	603	6 ABU11606	Abu11606 Human MDD
21	549.9	31.1	905	5 AAU76967	Aau76967 Novel hum
22	535.5	30.3	677	4 ABB65183	Abb65183 Drosophil
23	532.5	30.1	408	8 ADn40550	Adn40550 Candida a
24	532.5	30.1	408	8 ADN40554	Adn40554 Candida a
25	532.5	30.1	972	8 ADN40548	Adn40548 Candida a

ALIGNMENTS

RESULT 1

ADC23342

ID ADC23342 standard; protein; 346 AA.

XX AC ADC23342;

XX AC

DT 18-DEC-2003. (first entry)

XX DE Human kinesin-like DNA binding protein (KID) (SeqID 6).

XX KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID;

XX KW cytoskeletal; cardiac; immunomodulator; antiinflammatory; gene therapy;

XX KW cancer; hyperplasia; restenosis; cellular proliferation disorder;

XX KW cardiac hypertrophy; immune disorder; inflammation.

XX OS Homo sapiens.

XX PN US6387644-B1.

XX PD 14-MAY-2002.

XX PF 28-NOV-2000; 2000US-00724224.

XX PR 20-APR-1999; 99US-00295612.

XX PR 20-JUN-2000; 2000US-00597292.

XX XX (CYTO-) CYTOKINETICS INC.

XX Beraud C;

XX WPI; 2003-706919/67.

XX N-PSDB; ADC23341.

XX Identifying a candidate agent as modulator of function of a target

XX protein for treating cellular proliferation disorders by adding a

XX candidate agent to a mixture of the target protein that

XX directly/indirectly produces ADP or phosphate.

XX Claim 1; SEQ ID NO 6; 26pp; English.

XX This invention relates to a novel method for high throughput screening

XX systems used to identify compounds for the treatment of cellular

XX proliferation disorders. Specifically, it refers to candidate agents that

XX are capable of modulating the activity of target proteins having motor

XX domains, such that the target protein directly or indirectly produces ADP

XX or phosphate. Furthermore, this activity can be determined using a

XX fluorescence or absorbance readouts. The present invention describes a

XX method that identifies modulators of the target protein, which is a

Adn40552 Candida a
Abb07412 Amino aci
Abg72053 Human HsK
Abb07410 Human kin
Abg72052 Human kin
Aau19569 Human dia
Abp51294 Human MDD
Aae14401 Human HsK
Aae14402 Human HsK
Aau79592 Human HsK
Adc31540 Human nov
Aae14400 Human kin
Aau79590 Human kin
Abr48222 Human bla
Adb80468 Ovarian c
Adc35116 Human bre
Adl83290 Human PRO
Adq20128 Human sof
Adq09226 Human KNS
Adj95078 Novel NOV

CC kinesin-like DNA binding protein (known as KID) as cytoskeletal, cardinals,
 CC immunomodulators and antiinflammatory. Accordingly, through gene
 CC therapy, they can be used for the treatment of cancer, hyperplasia,
 CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This
 CC polypeptide sequence is human KID protein (SeqID 6) of the invention.
 XX
 SQ Sequence 346 AA;

Query Match 100.0%; Score 1768; DB 7; Length 346;
 Best Local Similarity 100.0%; Pred. No. 1.2e-173;
 Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGWDSCLSTANWRNHQ 60
 DB 1 MGRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGWDSCLSTANWRNHQ 60
 QY 61 TLKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYGPTGAGKTHMLGSPQP 120
 DB 61 TLKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYGPTGAGKTHMLGSPQP 120
 QY 121 GVIPALMDLLQLTREGAEGRPWALSVMTMSYLEIQKVLDDLLDPASGDLVIREDCRGN 180
 DB 121 GVIPALMDLLQLTREGAEGRPWALSVMTMSYLEIQKVLDDLLDPASGDLVIREDCRGN 180
 QY 181 ILIPGLSOKPISSPADFERHFLPASRNRTVGATRLNQRSSSHAVLLVKVDQERLAPFR 240
 DB 181 ILIPGLSOKPISSPADFERHFLPASRNRTVGATRLNQRSSSHAVLLVKVDQERLAPFR 240
 QY 241 QREGKLYLIDLAGESEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNQGLPRVPRYRDSK 300
 DB 241 QREGKLYLIDLAGESEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNQGLPRVPRYRDSK 300
 QY 301 LTRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
 DB 301 LTRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

RESULT 2
 ID ADO60232
 AC ADO60232 standard; protein; 346 AA.
 XX ADO60232;
 XX
 DT 23-SEP-2004 (first entry)
 DE Human microtubule motor protein #3.
 XX
 KW Human; microtubule motor protein; cellular proliferation disorder;
 KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;
 KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;
 KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.
 XX
 OS Homo sapiens.
 XX
 PN US6762043-B1.
 XX
 PD 13-JUL-2004.
 XX
 PF 06-MAR-2002; 2002US-000933317.
 XX
 PR 20-APR-1999; 99US-00295612.
 PR 20-JUN-2000; 2000US-00597292.
 PR 28-NOV-2000; 2000US-00724224.
 XX
 PA (CYTO-) CYTOKINETICS INC.
 XX
 PI Beraud C;
 XX
 XX WPI; 2004-532491/51.
 DR N-PSDB; ADO60231.
 XX
 PT New isolated microtubule motor protein, useful for screening modulators
 PT for treating cellular proliferation disorders such as cancer,

PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
 PT inflammation.
 XX
 PS Claim 1; SEQ ID NO 6; 26pp; English.
 XX
 CC The invention relates to human microtubule motor proteins and the nucleic
 CC acids encoding them. The invention also relates to a method of screening
 CC for modulators of a motor protein which has microtubule stimulated ATPase
 CC activity, a method of testing for ATPase activity of microtubule motor
 CC proteins, methods to identify candidate agents that bind to a target
 CC protein or act as a modulator of the binding characteristics or
 CC biological activity of a target protein, modulators of the target
 CC protein, and methods of treating cellular proliferation disorders such as
 CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
 CC and inflammation, for treating disorders associated with kinesin-like DNA
 CC binding protein (KID) and for inhibiting KID. The sequences are used for
 CC screening for modulators of motor proteins useful for treating cellular
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
 CC hypertrophy, immune disorders and inflammation, for treating disorders
 CC associated with KID and for inhibiting KID and for treating autoimmune
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and
 CC proliferation induced after medical procedures including surgery and
 CC angioplasty. This sequence represents a human microtubule motor protein
 CC of the invention.
 XX
 SQ Sequence 346 AA;

Query Match 100.0%; Score 1768; DB 8; Length 346;
 Best Local Similarity 100.0%; Pred. No. 1.2e-173;
 Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGWDSCLSTANWRNHQ 60
 DB 1 MGRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGWDSCLSTANWRNHQ 60
 QY 61 TLKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYGPTGAGKTHMLGSPQP 120
 DB 61 TLKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYGPTGAGKTHMLGSPQP 120
 QY 121 GVIPALMDLLQLTREGAEGRPWALSVMTMSYLEIQKVLDDLLDPASGDLVIREDCRGN 180
 DB 121 GVIPALMDLLQLTREGAEGRPWALSVMTMSYLEIQKVLDDLLDPASGDLVIREDCRGN 180
 QY 181 ILIPGLSOKPISSPADFERHFLPASRNRTVGATRLNQRSSSHAVLLVKVDQERLAPFR 240
 DB 181 ILIPGLSOKPISSPADFERHFLPASRNRTVGATRLNQRSSSHAVLLVKVDQERLAPFR 240
 QY 241 QREGKLYLIDLAGESEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNQGLPRVPRYRDSK 300
 DB 241 QREGKLYLIDLAGESEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNQGLPRVPRYRDSK 300
 QY 301 LTRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
 DB 301 LTRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

RESULT 3
 ID ADC23344
 AC ADC23344 standard; protein; 487 AA.
 XX ADC23344;
 XX
 DT 18-DEC-2003 (first entry)
 DE Human kinesin-like DNA binding protein (KID) (SeqID 8).
 XX
 KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID;
 KW cytoskeletal; cardinals; immunomodulator; antiinflammatory; gene therapy;
 KW cancer; hyperplasia; restenosis; cellular proliferation disorder;
 KW cardiac hypertrophy; immune disorder; inflammation.
 XX
 OS Homo sapiens.
 XX

PN US6387644-B1.
 XX 14-MAY-2002.
 XX 28-NOV-2000; 2000US-00724224.
 XX 20-APR-1999; 99US-00295612.
 XX 20-JUN-2000; 2000US-00597292.
 XX (CYTO-) CYTOKINETICS INC.
 XX Beraud C;
 XX WPI; 2003-706919/67.
 XX N-PSDB; ADC23343.
 XX Identifying a candidate agent as modulator of function of a target
 PT protein for treating cellular proliferation disorders by adding a
 PT candidate agent to a mixture of the target protein that
 PT directly/indirectly produces ADP or phosphate.
 XX Claim 1; SEQ ID NO 8; 26pp; English.
 XX This invention relates to a novel method for high throughput screening
 CC systems used to identify compounds for the treatment of cellular
 CC proliferation disorders. Specifically, it refers to candidate agents that
 CC are capable of modulating the activity of target proteins having motor
 CC domains, such that the target protein directly or indirectly produces ADP
 CC or phosphate. Furthermore, this activity can be determined using
 CC fluorescence or absorbance readouts. The present invention describes a
 CC method that identifies modulators of the target protein, which is a
 CC kinesin-like DNA binding protein (known as KID) as cytoskeletal, cardinals,
 CC immunomodulators and antiinflammatory. Accordingly, through gene
 CC therapy, they can be used for the treatment of cancer, hyperplasias,
 CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This
 CC polypeptide sequence is human KID protein (SeqID 8) of the invention.
 XX
 SQ Sequence 487 AA;
 Query Match 100.0%; Score 1768; DB 7; Length 487;
 Best Local Similarity 100.0%; Pred. No. 2.1e-173;
 Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGRCLSKIGATRRPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWNHQE 60
 Db 1 MGRCLSKIGATRRPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWNHQE 60
 QY 61 TLKYQDFAFYGERSTQDDIYAGSVQPIRLHLEGGQNASVLAIGPTGAGKTHMLGSPEQ 120
 Db 61 TLKYQDFAFYGERSTQDDIYAGSVQPIRLHLEGGQNASVLAIGPTGAGKTHMLGSPEQ 120
 QY 121 GVIPRALMDLLQITREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRG 180
 Db 121 GVIPRALMDLLQITREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRG 180
 QY 181 ILIPGLSQPISSFADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDORERLAPFR 240
 Db 181 ILIPGLSQPISSFADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDORERLAPFR 240
 QY 241 QREGKLYLIDLAGESSNRRGKGLKESGAINTSFLVLGKVVVDALNQLPRVPYRDSK 300
 Db 241 QREGKLYLIDLAGESSNRRGKGLKESGAINTSFLVLGKVVVDALNQLPRVPYRDSK 300
 QY 301 LTRLLQDSLGGSAHSILIANIAPERRFYLDVTVSALNFAARSKVIN 346
 Db 301 LTRLLQDSLGGSAHSILIANIAPERRFYLDVTVSALNFAARSKVIN 346
 RESULT 4
 ADQ60234 standard; protein; 487 AA.
 XX
 AC ADQ60234;

XX 23-SEP-2004 (first entry)
 XX Human microtubule motor protein #4.
 XX Human; microtubule motor protein; cellular proliferation disorder;
 KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;
 KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;
 KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.
 XX
 OS Homo sapiens.
 XX US6762043-B1.
 XX 13-JUL-2004.
 XX 06-MAR-2002; 2002US-00093317.
 XX 20-APR-1999; 99US-00295612.
 XX 20-JUN-2000; 2000US-00597292.
 XX 28-NOV-2000; 2000US-00724224.
 XX (CYTO-) CYTOKINETICS INC.
 XX Beraud C;
 XX WPI; 2004-532491/51.
 XX N-PSDB; ADQ60233.
 XX New isolated microtubule motor protein, useful for screening modulators
 PT for treating cellular proliferation disorders such as cancer,
 PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
 PT inflammation.
 XX Claim 1; SEQ ID NO 8; 26pp; English.
 XX The invention relates to human microtubule motor proteins and the nucleic
 CC acids encoding them. The invention also relates to a method of screening
 CC for modulators of a motor protein which has microtubule stimulated ATPase
 CC activity, a method of testing for ATPase activity of microtubule motor
 CC proteins, methods to identify candidate agents that bind to a target
 CC protein or act as a modulator of the binding characteristics or
 CC biological activity of a target protein, modulators of the target
 CC protein, and methods of treating cellular proliferation disorders such as
 CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
 CC and inflammation, for treating disorders associated with kinesin-like DNA
 CC binding protein (KID) and for inhibiting KID. The sequences are used for
 CC screening for modulators of motor proteins useful for treating cellular
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
 CC hypertrophy, immune disorders and inflammation, for treating disorders
 CC associated with KID and for inhibiting KID and for treating autoimmune
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and
 CC proliferation induced after medical procedures including surgery and
 CC angioplasty. This sequence represents a human microtubule motor protein
 CC of the invention.
 XX
 SQ Sequence 487 AA;
 Query Match 100.0%; Score 1768; DB 8; Length 487;
 Best Local Similarity 100.0%; Pred. No. 2.1e-173;
 Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGRCLSKIGATRRPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWNHQE 60
 Db 1 MGRCLSKIGATRRPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWNHQE 60
 QY 61 TLKYQDFAFYGERSTQDDIYAGSVQPIRLHLEGGQNASVLAIGPTGAGKTHMLGSPEQ 120
 Db 61 TLKYQDFAFYGERSTQDDIYAGSVQPIRLHLEGGQNASVLAIGPTGAGKTHMLGSPEQ 120
 QY 121 GVIPRALMDLLQITREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRG 180
 Db 121 GVIPRALMDLLQITREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRG 180

QY 181 ILIPGLSQPISSFADFERHFLPASNRVTGATRLNQRSSSRSHAVLLVKVDQERLAPR 240
 DB 181 ILIPGLSQPISSFADFERHFLPASNRVTGATRLNQRSSSRSHAVLLVKVDQERLAPR 240
 QY 241 QREGKLYLIDLAGSDNRTGNKGLRKESGAINSTSLFVLGKVDALNOGLPRVPYRDSK 300
 DB 241 QREGKLYLIDLAGSDNRTGNKGLRKESGAINSTSLFVLGKVDALNOGLPRVPYRDSK 300
 QY 301 LTRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
 DB 301 LTRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

RESULT 5
 ADC23338
 ID ADC23338 standard; protein; 370 AA.

XX AC ADC23338;

XX DT 18-DEC-2003 (first entry)

XX DE Human kinesin-like DNA binding protein (KID) (SeqID 2).

XX KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID;
 KW cytoskeletal; cardiac; immunomodulator; antiinflammatory; gene therapy;
 KW cancer; hyperplasia; restenosis; cellular proliferation disorder;
 KW cardiac hypertrophy; immune disorder; inflammation.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Misc-difference 1 /note= "Encoded by ATGCA"

XX PN US6387644-B1.

XX PD 14-MAY-2002.

XX PF 28-NOV-2000; 2000US-00724224.

XX PR 20-APR-1999; 99US-00295612.

XX PR 20-JUN-2000; 2000US-00597292.

XX PA (CYTO-) CYTOKINETICS INC.

XX PI Beraud C;

XX WPI; 2003-706919/67.

XX DR N-PSDB; ADC23337.

XX PT Identifying a candidate agent as modulator of function of a target
 PT protein for treating cellular proliferation disorders by adding a
 PT candidate agent to a mixture of the target protein that
 PT directly/indirectly produces ADP or phosphate.

XX PS Claim 1; SEQ ID NO 2; 26pp; English.

XX CC This invention relates to a novel method for high throughput screening
 CC systems used to identify compounds for the treatment of cellular
 CC proliferation disorders. Specifically, it refers to candidate agents that
 CC are capable of modulating the activity of target proteins having motor
 CC domains, such that the target protein directly or indirectly produces ADP
 CC or phosphate. Furthermore, this activity can be determined using
 CC fluorescence or absorbance readouts. The present invention describes a
 CC method that identifies modulators of the target protein, which is a
 CC kinesin-like DNA binding protein (known as KID) as cytoskeletal, cardiac,
 CC immunomodulators and antiinflammatory. Accordingly, through gene
 CC therapy, they can be used for the treatment of cancer, hyperplasias,
 CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This
 CC polypeptide sequence is human KID protein (SeqID 2) of the invention.

XX SQ Sequence 370 AA;

Query Match 99.7%; Score 1763; DB 7; Length 370;
 Best Local Similarity 100.0%; Pred. No. 4.5e-173; Indels 0; Gaps 0;
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GRCLSKIGATRRPPPARVRVAVRLRPFVDTAGAGDPPPCVRGMDSCSLEIANRNHQET 61
 DB 26 GRCLSKIGATRRPPPARVRVAVRLRPFVDTAGAGDPPPCVRGMDSCSLEIANRNHQET 85
 QY 62 LKYOFAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAIGPTGAGKTHMLGSPPOPG 121
 DB 86 LKYOFAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAIGPTGAGKTHMLGSPPOPG 145
 QY 122 VIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDDLLDPASGDLVIREDCRGN 181
 DB 146 VIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDDLLDPASGDLVIREDCRGN 205
 QY 182 LIPGLSQPISSFADFERHFLPASNRVTGATRLNQRSSSRSHAVLLVKVDQERLAPRQ 241
 DB 206 LIPGLSQPISSFADFERHFLPASNRVTGATRLNQRSSSRSHAVLLVKVDQERLAPRQ 265
 QY 242 REGKLYLIDLAGSDNRTGNKGLRKESGAINSTSLFVLGKVDALNOGLPRVPYRDSKL 301
 DB 266 REGKLYLIDLAGSDNRTGNKGLRKESGAINSTSLFVLGKVDALNOGLPRVPYRDSKL 325
 QY 302 LTRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
 DB 326 LTRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 370

RESULT 6

ID ADQ60228 standard; protein; 370 AA.

XX AC ADQ60228;

XX DT 23-SEP-2004 (first entry)

XX DE Human microtubule motor protein #1.

XX KW Human; microtubule motor protein; cellular proliferation disorder;
 KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;
 KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;
 KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.

XX OS Homo sapiens.

XX PN US6762043-B1.

XX PD 13-JUL-2004.

XX PF 06-MAR-2002; 2002US-00093317.

XX PR 20-APR-1999; 99US-00295612.

XX PR 20-JUN-2000; 2000US-00597292.

XX PR 28-NOV-2000; 2000US-00724224.

XX PA (CYTO-) CYTOKINETICS INC.

XX PI Beraud C;

XX WPI; 2004-532491/51.

XX PT New isolated microtubule motor protein, useful for screening modulators
 PT for treating cellular proliferation disorders such as cancer,
 PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
 PT inflammation.

XX PS Claim 1; SEQ ID NO 2; 26pp; English.

XX CC The invention relates to human microtubule motor proteins and the nucleic
 CC acids encoding them. The invention also relates to a method of screening
 CC for modulators of a motor protein which has microtubule stimulated ATPase

CC activity, a method of testing for ATPase activity of microtubule motor
 CC proteins, methods to identify candidate agents that bind to a target
 CC protein or act as a modulator of the binding characteristics or
 CC biological activity of a target protein, modulators of the target
 CC protein, and methods of treating cellular proliferation disorders such as
 CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
 CC and inflammation, for treating disorders associated with kinesin-like DNA
 CC binding protein (KID) and for inhibiting KID. The sequences are used for
 CC screening for modulators of motor proteins useful for treating cellular
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
 CC hypertrophy, immune disorders and inflammation, for treating disorders
 CC associated with KID and for inhibiting KID and for treating autoimmune
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and
 CC proliferation induced after medical procedures including surgery and
 CC angioplasty. This sequence represents a human microtubule motor protein
 CC of the invention. Note: The specification states that this sequence is
 CC encoded by the nucleic acid featured as SEQ ID NO:1, but this does not
 CC appear to be the case.

XX SQ Sequence 370 AA;

Query Match 99.7%; Score 1763; DB 8; Length 370;
 Best Local Similarity 100.0%; Pred. No. 4.5e-173;
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVRVAVRLRPFDVGTAGASDPPCVRGMDSCSLEIANWRNHQET 61
 DB 26 GRCLSKIGATRRPPPARVRVAVRLRPFDVGTAGASDPPCVRGMDSCSLEIANWRNHQET 85
 QY 62 LKYPDAFYGERSTQODIYAGSVQPIRLHLEQNASVLA YGPTGAGKTHMLGSPQPG 121
 DB 86 LKYPDAFYGERSTQODIYAGSVQPIRLHLEQNASVLA YGPTGAGKTHMLGSPQPG 145
 QY 122 VIPRALMDLQLLTREEGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 191
 DB 146 VIPRALMDLQLLTREEGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 205
 QY 182 LIPGLSQKPISSFADFERHFLPASRNRVTGATRLNQRSSSHAVLVKVDQERLAPFRQ 241
 DB 206 LIPGLSQKPISSFADFERHFLPASRNRVTGATRLNQRSSSHAVLVKVDQERLAPFRQ 265
 QY 242 REGKLYLIDLAGSEDRRTGNKGLRKESGAINTSFLVGLKVVDALNQGLPRVPYRDSKL 301
 DB 266 REGKLYLIDLAGSEDRRTGNKGLRKESGAINTSFLVGLKVVDALNQGLPRVPYRDSKL 325
 QY 302 TRLLQDSLGGSAHSILIANIAPERRFYLDITVSALNFAARKEVIN 346
 DB 326 TRLLQDSLGGSAHSILIANIAPERRFYLDITVSALNFAARKEVIN 370

RESULT 7

ID ADK40973 standard; protein; 490 AA.

XX AC ADK40973;

XX DT 06-MAY-2004 (first entry)

XX DE Novel human kinase protein #80.

XX cytosolic; immunomodulator; cardiant; neuroprotective; nootropic;
 KW antiparkinsonian; virucide; antibacterial; fungicide; ophthalmological;
 KW analgesic; hypotensive; immunosuppressive; kinase inhibitor; kinase;
 KW cancer; peripheral nervous system; central nervous system;
 KW Alzheimer's disease; Parkinson's disease; multiple sclerosis;
 KW amyotrophic lateral sclerosis; viral infection; prion infection;
 KW ocular disease; migraine; pain; sexual dysfunction; mood disorder;
 KW attention disorder; cognition disorder; hypotension; hypertension;
 KW psychotic disorder; neurological disorder; dyskinesia;
 KW metabolic disorder; organ transplant rejection; enzyme.

XX Homo sapiens.

OS

PN WO2003057841-A2.
 XX 17-JUL-2003.
 PD
 XX
 XX
 PF 31-DEC-2002; 2002WO-US041687.
 XX
 PR 31-DEC-2001; 2001US-0343169P.
 XX
 PA (GRIG/) GRIGORIEV I V.
 PA (SUDA/) SUDARSANAM S.
 XX
 XX Grigoriev IV, Sudarsanam S;
 PI
 XX WPI; 2003-587115/55.
 DR
 XX
 XX New isolated, enriched or purified nucleic acid molecule encoding a
 PT kinase polypeptide, useful for treating cancer, immune-related diseases,
 PT cardiovascular disease, brain or neuronal-associated diseases and
 PT metabolic disorders.
 XX
 XX
 PS Claim 1; SEQ ID NO 80; 491pp; English.

CC The invention relates to novel isolated, enriched or purified nucleic acid
 CC molecules encoding a kinase polypeptide. The nucleic acid molecule
 CC comprises a sequence that: (a) encodes a kinase polypeptide; (b) is a
 CC complement of (a); (c) hybridizes under stringent conditions to (a) and
 CC encodes a naturally occurring kinase polypeptide; (d) encodes the
 CC polypeptide in (a), except that it lacks one or more, but not all, of an
 CC N-terminal domain, C-terminal catalytic domain, a catalytic domain, a C-
 CC terminal domain, a coiled-coil structure region, a spacer region and a C-
 CC terminal tail; or (e) is a complement of (d). The nucleic acid molecules,
 CC polypeptides, methods and substance are useful for treating cancers,
 CC immune-related diseases or disorders, cardiovascular disease, brain or
 CC neuronal-associated diseases, and metabolic disorders. The disorders are
 CC preferably cancers of the tissues or of hematopoietic origin, diseases of
 CC the central or peripheral nervous system, Alzheimer's disease,
 CC Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis,
 CC viral infections, infections caused by prions, infections caused by
 CC bacteria, infections caused by fungi, ocular diseases, migraines, pain,
 CC sexual dysfunction, mood disorders, attention disorders, cognition
 CC disorders, hypotension, hypertension, psychotic disorders, neurological
 CC disorders, dyskinesias, metabolic disorders and organ transplant
 CC rejection. This sequence corresponds to one of the kinase polypeptides of
 CC the invention.

XX SQ Sequence 490 AA;

Query Match 99.7%; Score 1763; DB 7; Length 490;
 Best Local Similarity 100.0%; Pred. No. 6.9e-173;
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVRVAVRLRPFDVGTAGASDPPCVRGMDSCSLEIANWRNHQET 61
 DB 34 GRCLSKIGATRRPPPARVRVAVRLRPFDVGTAGASDPPCVRGMDSCSLEIANWRNHQET 93
 QY 62 LKYPDAFYGERSTQODIYAGSVQPIRLHLEQNASVLA YGPTGAGKTHMLGSPQPG 121
 DB 94 LKYPDAFYGERSTQODIYAGSVQPIRLHLEQNASVLA YGPTGAGKTHMLGSPQPG 153
 QY 122 VIPRALMDLQLLTREEGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 181
 DB 154 VIPRALMDLQLLTREEGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 213
 QY 182 LIPGLSQKPISSFADFERHFLPASRNRVTGATRLNQRSSSHAVLVKVDQERLAPFRQ 241
 DB 214 LIPGLSQKPISSFADFERHFLPASRNRVTGATRLNQRSSSHAVLVKVDQERLAPFRQ 273
 QY 242 REGKLYLIDLAGSEDRRTGNKGLRKESGAINTSFLVGLKVVDALNQGLPRVPYRDSKL 301
 DB 274 REGKLYLIDLAGSEDRRTGNKGLRKESGAINTSFLVGLKVVDALNQGLPRVPYRDSKL 333
 QY 302 TRLLQDSLGGSAHSILIANIAPERRFYLDITVSALNFAARKEVIN 346

Db 334 TRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 378

RESULT 8
ADC23340
ID ADC23340 standard; protein; 512 AA.
XX
AC ADC23340;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human kinesin-like DNA binding protein (KID) (SeqID 4).
XX
KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID;
KW cytosolic; cardiant; immunomodulator; antiinflammatory; gene therapy;
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;
KW cardiac hypertrophy; immune disorder; inflammation.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 2 /note= "Encoded by CA"
FT
XX
FN US6387644-B1.
PD 14-MAY-2002.
XX
PF 28-NOV-2000; 2000US-00724224.
XX
PR 20-APR-1999; 99US-00295612.
PR 20-JUN-2000; 2000US-00597292.
XX
PA (CYTO-) CYTOKINETICS INC.
XX
PI Beraud C;
XX
DR WPI; 2003-706919/67.
DR N-PSDB; ADC23339.
XX
XX Identifying a candidate agent as modulator of function of a target
PT protein for treating cellular proliferation disorders by adding a
PT candidate agent to a mixture of the target protein that
PT directly/indirectly produces ADP or phosphate.
XX
PS Claim 1; SEQ ID NO 4; 26pp; English.
XX
CC This invention relates to a novel method for high throughput screening
CC systems used to identify compounds for the treatment of cellular
CC proliferation disorders. Specifically, it refers to candidate agents that
CC are capable of modulating the activity of target proteins having motor
CC domains, such that the target protein directly or indirectly produces ADP
CC or phosphate. Furthermore, this activity can be determined using
CC fluorescence or absorbance readouts. The present invention describes a
CC method that identifies modulators of the target protein, which is a
CC kinesin-like DNA binding protein (known as KID) as cytostatic, cardiants,
CC immunomodulators and antiinflammatories. Accordingly, through gene
CC therapy, they can be used for the treatment of cancer, hyperplasias,
CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This
CC polypeptide sequence is human KID protein (SeqID 4) of the invention.
XX
SQ Sequence 512 AA;
Query Match 99.7%; Score 1763; DB 7; Length 512;
Best Local Similarity 100.0%; Pred. No. 7.4e-173;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GRCRLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRCWMDSCSLEIANWRHQET 61
Db 27 GRCRLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRCWMDSCSLEIANWRHQET 86
QY 62 LKQFDIFYGERSTQDDIYAGSVQPIIRHLLEGQNASVLAYGPTGAGKTHMLGSPQPG 121

Db 87 LKQFDIFYGERSTQDDIYAGSVQPIIRHLLEGQNASVLAYGPTGAGKTHMLGSPQPG 146

QY 122 VIPRALMDLLQLTREGAERPWALSVTMSYLIYQEKVLDLDPASGDLVIREDCRNI 181
147 VIPRALMDLLQLTREGAERPWALSVTMSYLIYQEKVLDLDPASGDLVIREDCRNI 206

QY 182 LIPGLSOKPISSPADFERHFLPASRNRVTGATRLNQRSSSRSHAVLLVKVQERELAPFRQ 241

Db 207 LIPGLSOKPISSPADFERHFLPASRNRVTGATRLNQRSSSRSHAVLLVKVQERELAPFRQ 266

QY 242 REGKYLIDLAGSEDNRRGTGKGLRKESGAINTSLFVLGKVVQDALNOGLPRVYRDSKL 301

Db 267 REGKYLIDLAGSEDNRRGTGKGLRKESGAINTSLFVLGKVVQDALNOGLPRVYRDSKL 326

QY 302 TRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

Db 327 TRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 371

RESULT 9
ADQ60230
ID ADQ60230 standard; protein; 512 AA.
XX
AC ADQ60230;
XX
DT 23-SEP-2004 (first entry)
XX
DE Human microtubule motor protein #2.
XX
KW Human; microtubule motor protein; cellular proliferation disorder;
KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;
KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;
KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.
XX
OS Homo sapiens.
XX
PN US6762043-B1.
PD 13-JUL-2004.
XX
PF 06-MAR-2002; 2002US-00093317.
XX
PR 20-APR-1999; 99US-00295612.
PR 20-JUN-2000; 2000US-00597292.
PR 28-NOV-2000; 2000US-00724224.
XX
PA (CYTO-) CYTOKINETICS INC.
XX
PI Beraud C;
XX
DR WPI; 2004-532491/51.
XX
PT New isolated microtubule motor protein, useful for screening modulators
PT for treating cellular proliferation disorders such as cancer,
PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
PT inflammation.
XX
PS Claim 1; SEQ ID NO 4; 26pp; English.
XX
CC The invention relates to human microtubule motor proteins and the nucleic
CC acids encoding them. The invention also relates to a method of screening
CC for modulators of a motor protein which has microtubule stimulated ATPase
CC activity, a method of testing for ATPase activity of microtubule motor
CC proteins, methods to identify candidate agents that bind to a target
CC protein or act as a modulator of the binding characteristics of the target
CC biological activity of a target protein, modulators of the target
CC protein, and methods of treating cellular proliferation disorders such as
CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
CC and inflammation, for treating disorders associated with kinesin-like DNA
CC binding protein (KID) and for inhibiting KID. The sequences are used for
CC screening for modulators of motor proteins useful for treating cellular
CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
CC hypertrophy, immune disorders and inflammation, for treating disorders

CC associated with KID and for inhibiting KID and for treating autoimmune
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and
 CC proliferation induced after medical procedures including surgery and
 CC angioplasty. This sequence represents a human microtubule motor protein
 CC of the invention. Note: The specification states that this sequence is
 CC encoded by the nucleic acid featured as SEQ ID NO:3, but this does not
 CC appear to be the case.

XX
 SQ Sequence 512 AA;

Query Match 99.7%; Score 1763; DB 8; Length 512;
 Best Local Similarity 100.0%; Pred. No. 7.4e-173; Indels 0; Gaps 0;
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVRVAVLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61
 DB 27 GRCLSKIGATRRPPPARVRVAVLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 86

QY 62 LKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 121
 DB 87 LKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 146

QY 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 181
 DB 147 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 206

QY 182 LIPGLSQKPISSPADFERHFLPASRNRVTGATRLNQSRSSHAVLLVKVQDRLAPFRQ 241
 DB 207 LIPGLSQKPISSPADFERHFLPASRNRVTGATRLNQSRSSHAVLLVKVQDRLAPFRQ 266

QY 242 REGKLYLDLAGSDNRRTGNKGLRKESGAINTSFLVLGKVDVALNOGLPRVPYRDSKL 301
 DB 267 REGKLYLDLAGSDNRRTGNKGLRKESGAINTSFLVLGKVDVALNOGLPRVPYRDSKL 326

QY 302 TRLLQDSLGGSAAHSILIANIAPERFYLDTVSALNFAARKEVIN 346
 DB 327 TRLLQDSLGGSAAHSILIANIAPERFYLDTVSALNFAARKEVIN 371

RESULT 10

ADQ09240
 ID ADQ09240 standard; protein; 665 AA.

AC ADQ09240;

XX 23-SEP-2004 (first entry)

DE Human KNSL4 protein SEQ ID NO:425.

XX thanatos-associated protein; THAP; THAP responsive gene; THAP family;
 KW THAP responsive element; angiogenesis; inflammation; metastasis; cancer;
 KW apoptosis; cardiovascular disease; neurodegenerative disease; chemokine;
 KW antiangiogenic; antiinflammatory; cardiovascular; cytoskeletal;
 KW neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;
 KW human.

OS Homo sapiens.

XX WQ2004055050-A2.

PN 01-JUL-2004.

PD 10-DEC-2003; 2003MO-IB006434.

PF 10-DEC-2002; 2002US-0432699P.

PR 03-JUL-2003; 2003US-0485027P.

XX (ENDO-) ENDOCUBE SAS.

PA (CNRS) CNRS CENT NAT RECH SCI.

PI Girard J, Amalric F, Roussigne M, Clouaire T;

XX WPI; 2004-525034/50.

DR N-PSDB; ADQ09241.

XX Modulating expression of a Thanatos (death)-Associated Protein (THAP)
 PT responsive gene for preventing or treating e.g. cancer or inflammation,
 PT comprises modulating the interaction of a THAP polypeptide with a nucleic
 PT acid.

XX Example 47; SEQ ID NO 425; 612pp; English.

XX The present invention describes a method for modulating the expression of
 CC a thanatos (death)-associated protein (THAP) responsive gene. The method
 CC comprises modulating the interaction of a THAP-family polypeptide or its
 CC biological fragment with a nucleic acid, and so enhancing or repressing
 CC the expression of the THAP responsive gene. Also described: (1) a method
 CC of modulating the expression of a gene responsive to a THAP/chemokine
 CC complex; (2) a pharmaceutical composition comprising a THAP responsive
 CC element in a pharmaceutical carrier; (3) a transcription factor decoy
 CC consisting essentially of a THAP responsive element; (4) a cell
 CC comprising a transcription factor decoy described above; (5) methods of
 CC modulating the interaction between a nucleic acid and a THAP-family
 CC polypeptide or its biological fragment, or a nucleic acid and a
 CC THAP/chemokine complex; (6) a vector packaging cell line comprising a
 CC cell comprising a viral vector which comprises a promoter operably linked
 CC to a nucleic acid encoding a THAP-family polypeptide or its biological
 CC fragment; (7) a method of constructing a cell which expresses a
 CC recombinant THAP-family polypeptide; (8) a method of ameliorating
 CC symptoms associated with a condition mediated by a THAP/chemokine complex
 CC; (9) methods of identifying a test compound that modulates transcription
 CC at a THAP responsive element or that modulates the transport of a
 CC chemokine into the nucleus; (10) methods for reducing the symptoms
 CC associated with a condition selected from excessive or insufficient
 CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
 CC or insufficient apoptosis, cardiovascular disease and neurodegenerative
 CC diseases; symptoms associated with a condition resulting from the
 CC activity of a chemokine or a THAP-family polypeptide in an individual; or
 CC symptoms associated with transcriptional repression or activation
 CC mediated by a THAP-family polypeptide in an individual; (11) a vector
 CC comprising a THAP responsive promoter operably linked to a nucleic acid
 CC encoding a detectable product; (12) a genetically engineered cell
 CC comprising the vector described above, or that expresses a THAP-family
 CC polypeptide or its biological fragment; (13) an in vitro transcription
 CC reaction comprising a nucleic acid comprising a THAP responsive promoter,
 CC ribonucleotides and an RNA polymerase; and (14) an isolated mutant THAP-
 CC family polypeptide that does not bind to a chemokine. The pharmaceutical
 CC composition has antiangiogenic, antiinflammatory, cardiovascular,
 CC cytostatic, neuroprotective and osteopathic activities, and can be used
 CC as a THAP and THAP synthesis modulator. The composition can be used for
 CC modulating the expression of a THAP responsive gene. Modulation is useful
 CC for reducing symptoms of conditions such as excessive or insufficient
 CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
 CC or insufficient apoptosis, cardiovascular disease or neurodegenerative
 CC diseases. The present sequence is used in the exemplification of the
 CC present invention.

XX SQ Sequence 665 AA;

Query Match 99.7%; Score 1763; DB 8; Length 665;

Best Local Similarity 100.0%; Pred. No. 1.1e-172;

Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVRVAVLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61

DB 26 GRCLSKIGATRRPPPARVRVAVLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 85

QY 62 LKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 121

DB 86 LKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 145

QY 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 181

DB 146 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 205

QY 182 LIPGLSQKPISSPADFERHFLPASRNRVTGATRLNQSRSSHAVLLVKVQDRLAPFRQ 241

Db 206 LIPGLSQKPISSFADFERHFLPASRRTVGATRLNQRSSSHAVLLVKVDQRLAPFRQ 265
QY 242 REGKLYLIDLAGSEDRRTGKGLRKESGAINTSFLVLGKVVDAALNOGLPRVPYRDSKL 301
Db 266 REGKLYLIDLAGSEDRRTGKGLRKESGAINTSFLVLGKVVDAALNOGLPRVPYRDSKL 325
QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346
Db 326 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 370
RESULT 11
ID AAB56650 standard; protein; 460 AA.
XX AAB56650;
AC AAB56650;
XX 13-MAR-2001 (first entry)
XX Human prostate cancer antigen protein sequence SEQ ID NO:1228.
XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW vulnary; gastrointestinal; nephrotropic; anti-infective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
Wound; infectious disease.
XX Homo sapiens.
OS
XX WO200055174-A1.
PN
XX 21-SEP-2000.
PD
XX 08-MAR-2000; 2000WO-US005988.
PF
XX 12-MAR-1999; 99US-0124270P.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
PA
PI Rosen CA, Ruben SM;
XX WPI; 2000-587513/55.
XX N-PSDB; AAF15853.
DR
XX Prostate cancer associated gene sequences, referred to as prostate cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as prostate cancer.
XX
XX Claim 11; Page 1649-1651; 2338pp; English.
XX
XX AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens. Given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
CC nephrotropic, anti-infective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 460 AA;

Query Match 99.2%; Score 1753; DB 3; Length 460;
Best Local Similarity 99.4%; Pred. No. 6.8e-172;
Matches 343; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLETANRNHOET 61
Db 4 GRCLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLETANRNHOET 63
QY 62 LKYOFDAFYGBRSQOQDIYAGSVOPILRHLLLEGONASVLAYGPTGAGKTHTMLGSPQPG 121
Db 64 LKYOFDAFYGBRSQOQDIYAGSVOPILRHLLLEGONASVLAYGPTGAGKTHTMLGSPQPG 123
QY 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQKVLDDLDPASGDLVIREDCRNI 181
Db 124 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQKVLDDLDPASGDLVIREDCRNI 183
QY 182 LIPGLSQKPISSFADFERHFLPASRRTVGATRLNQRSSSHAVLLVKVDQRLAPFRQ 241
Db 184 LIPGLSQKPISSFADFERHFLPASRRTVGATRLNQRSSSHAVLLVKVDQRLAPFRQ 243
QY 242 REGKLYLIDLAGSEDRRTGKGLRKESGAINTSFLVLGKVVDAALNOGLPRVPYRDSKL 301
Db 244 REGKLYLIDLAGSEDRRTGKGLRKESGAINTSFLVLGKVVDAALNOGLPRVPYRDSKL 303
QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346
Db 304 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 348
RESULT 12
ID AAB71112 standard; protein; 784 AA.
XX AAB71112;
AC AAB71112;
XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster polypeptide SEQ ID NO 40128.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
KW
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
XX PN
XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001WO-US009231.
PF
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
XX N-PSDB; ABL15215.
DR
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 40128; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABLO1840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

```

XX SQ Sequence 784 AA;
Query Match 31.6%; Score 559.5; DB 4; Length 784;
Best Local Similarity 39.8%; Pred. No. 4.4e-48;
Matches 141; Conservative 56; Mismatches 126; Indels 31; Gaps 10;

QY 12 TRRP-----PPARVAVRLRPFVDGTAGASDPPCV-----RGMDSCSLEIANW--RN 57
D 11 QLMVALRVRPISVALEEGATLAHKVDEQWVLMMPMEDPDDILRAHRSREKSYLFDA 70
QY 58 HOETLKQVDFYGERSTQDDIYAGSVQPIRLHLEGGONASVLAEGTGTGAGKTHM---L 114
D 61 KEQRKFTYDAAYDASATQTTLYHEVVPFLVSSVLEGFCIFAYGQTGTGKTFTMEGVR 120
QY 115 GSPQPGVIRPALMDL-LQLTREBGAEGRPWALSVTMSYLEIYQEKVLDDLPASGDLVI 173
D 121 GNDELMIIPRTFQIWLHNTEN-----FOFLVDVSYLEIYMEELRDLKPNKSHLEV 175
QY 174 REDCRGN-ILIPGLSQKPISSFADFERHFLPASRNRTVGATRLNQRSSRSRSHAVLLVKVDQ 232
D 176 RE--RSGGVVYVNLHAINCKSVEDMIKVMQVGNKRTVGFTNMVHSSRSRSHAFIMKIEM 233
QY 233 RERLAPRQRECKLYLIDLAGSEDRNRRTGKGLRKESGAINSTLFLVGLKVVDMALNOGLP 292
D 234 CD-TETNTIKVGLKLNLDLAGSERQSKTGASAEKRLKEASKINLALSLGNVISALAEESP 292
QY 293 RVPYRDSKRLRLQDSLGSAHSILIANIAPERFRFYLDVTSALNFAARSKEVIN 346
D 293 HVPYRDSKRLRLQDSLGSKTIIMIANIGPSNNTYNETUTLTIYASRAKSION 346

RESULT 13
ADM04007
ID ADM04007 standard; protein; 548 AA.
AC ADM04007;
XX
XX 20-MAY-2004 (first entry)
XX
XX Human protein of the invention SEQ ID NO:2692.
XX
XX human; gene therapy; diagnostic marker; pharmaceutical.
XX
XX Homo sapiens.
XX
XX EP1347046-A1.
XX
XX 24-SEP-2003.
XX
XX 12-APR-2002; 2002EP-00008400.
XX
XX 22-MAR-2002; 2002JP-00137785.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
XX WPI; 2003-723558/69.
XX
XX N-PSDB; ADM01564.
XX
XX New polynucleotides and polypeptides are useful in gene therapy, for
PT developing a diagnostic marker or medicines for regulating their
PT expression and activity, or as a target of gene therapy.
XX
XX Claim 1; SEQ ID NO 2692; 305pp; English.
XX
XX The invention relates to a novel human polynucleotide and the encoded
CC polypeptide. A polynucleotide of the invention may have a use in gene
CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
CC as a primer for synthesizing the polynucleotide or as a probe for

```

```

CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
CC useful in gene therapy, for developing a diagnostic marker or medicines
CC for regulating their expression and activity, or as a target of gene
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
CC are useful as pharmaceutical agents. The present sequence represents a
CC protein sequence of the invention.
XX
XX SQ Sequence 548 AA;
Query Match 31.6%; Score 559; DB 7; Length 548;
Best Local Similarity 38.9%; Pred. No. 2.8e-48;
Matches 132; Conservative 70; Mismatches 121; Indels 16; Gaps 7;

QY 19 RRVAVRLRPF--VDGTAGAS-----DPPCVRGMDSCSLEIANWRNHQETLK-YQDFAF 69
D 11 QLMVALRVRPISVALEEGATLAHKVDEQWVLMMPMEDPDDILRAHRSREKSYLFDA 70
QY 70 YGERSQDDIYAGSVQPIRLHLEGGONASVLAEGTGTGAGKTHMLGSPQPGVIRPALMD 129
D 71 FDFATQEMVYQATTKSLIEGVISGVNATVFAVGPTGCGKTYTLMGTQBPQIVVQTLND 130
QY 130 LLQLTREBGAEGRPWALSVTMSYLEIYQEKVLDDLPASGDLVIREDCRGNILIPGLSOK 189
D 131 LFRABETSND---MEYEVSMYSLEYNEIMIRLLNPSGLYLELREDSKGVIVAGITEV 187
QY 190 PISFADFERHFLPASRNRTVGATRLNQRSSRSRSHAVLLVKVDORERLAPFRQ--REGKLY 247
D 188 STINAKEMQLLMKGNRQRTQPTAANQTSRSRSHAVLQVTVRQSRVKNILQEVQRGLP 247
QY 248 LIDLAGSEDRNRRTGKGLRKESGAINSTLFLVGLKVVDAI-NOGLPR-VPYRDSKRLRL 305
D 248 MIDLAGSERASQTONRGQRMKEGAINRSLALGNCINALSDKGNKYNIVRDSKLTRELL 307
QY 306 QDSLGSAHSILIANIAPERFRFYLDVTSALNFAARSKEV 344
D 308 KDSLGNSRTVMTAHISPASSAFEESRNTLTLYAGRAKNI 346

RESULT 14
ADM04167
ID ADM04167 standard; protein; 998 AA.
XX
XX ADM04167;
XX
XX 15-JUL-2004 (first entry)
XX
XX Structural and cytoskeleton-associated polypeptide #3.
XX
XX structural and cytoskeleton-associated polypeptide; SCAP;
KW cell proliferative disorder; cancer; atherosclerosis; viral infection;
KW HIV; neurological disorder; Parkinson's disease; Alzheimer's disease;
KW stroke; human.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Domain 9
FT Domain /note = kinesin motor catalytic domain
FT Domain 11..25
FT Modified-site /note = kinesin motor domain
FT Modified-site 22
FT Domain /note= "potential phosphorylation site"
FT Domain 62..77
FT Domain /note = kinesin motor domain
FT Domain 63..647
FT Modified-site /note = kinesin motor domain
FT Modified-site 65
FT Modified-site /note= "potential phosphorylation site"
FT Modified-site 84
FT Modified-site /note= "potential phosphorylation site"
FT Modified-site 87
FT Domain /note= "potential phosphorylation site"
FT Domain 95..116

```

FT Modified-site /note = kinesin motor domain
 FT 97
 FT Modified-site /note = "potential glycosylation site"
 FT 117
 FT Modified-site /note = "potential phosphorylation site"
 FT 122
 FT Domain /note = kinesin motor domain
 FT 127
 FT Modified-site /note = "potential phosphorylation site"
 FT 138
 FT Modified-site /note = "potential phosphorylation site"
 FT 144
 FT Domain /note = kinesin motor domain
 FT 149
 FT Modified-site /note = "potential phosphorylation site"
 FT 202
 FT Domain /note = kinesin motor domain
 FT 214
 FT Modified-site /note = "potential glycosylation site"
 FT 217
 FT Modified-site /note = "potential phosphorylation site"
 FT 227
 FT Modified-site /note = "potential phosphorylation site"
 FT 248
 FT Domain /note = kinesin motor domain
 FT 254
 FT Modified-site /note = "potential phosphorylation site"
 FT 274
 FT Modified-site /note = "potential glycosylation site"
 FT 288
 FT Modified-site /note = "potential phosphorylation site"
 FT 292
 FT Modified-site /note = "potential phosphorylation site"
 FT 295
 FT Domain /note = kinesin motor domain
 FT 328
 FT Modified-site /note = "potential phosphorylation site"
 FT 356
 FT Modified-site /note = "potential glycosylation site"
 FT 386
 FT Modified-site /note = "potential phosphorylation site"
 FT 397
 FT Region /note = "leucine zipper pattern"
 FT 454
 FT Modified-site /note = "potential phosphorylation site"
 FT 469
 FT Modified-site /note = "potential phosphorylation site"
 FT 489
 FT Modified-site /note = "potential phosphorylation site"
 FT 495
 FT Modified-site /note = "potential phosphorylation site"
 FT 518
 FT Region /note = "leucine zipper pattern"
 FT 525
 FT Region /note = "leucine zipper pattern"
 FT 605
 FT Modified-site /note = "potential phosphorylation site"
 FT 649
 FT Modified-site /note = "potential phosphorylation site"
 FT 678
 FT Modified-site /note = "potential phosphorylation site"
 FT 682
 FT Modified-site /note = "potential phosphorylation site"
 FT 703
 FT Modified-site /note = "potential phosphorylation site"
 FT 753
 FT Modified-site /note = "potential phosphorylation site"
 FT 760
 FT Modified-site /note = "potential glycosylation site"
 FT 767
 FT Modified-site /note = "potential phosphorylation site"
 FT 790
 FT Modified-site /note = "potential phosphorylation site"

FT Modified-site 805
 FT /note = "potential phosphorylation site"
 FT 814
 FT Modified-site /note = "potential phosphorylation site"
 FT 816
 FT Modified-site /note = "potential phosphorylation site"
 FT 845
 FT Modified-site /note = "potential glycosylation site"
 FT 849
 FT Modified-site /note = "potential phosphorylation site"
 FT 893
 FT Modified-site /note = "potential phosphorylation site"
 FT 973
 FT Modified-site /note = "potential phosphorylation site"
 FT 977
 FT Modified-site /note = "potential phosphorylation site"

WO2004029205-A2.

08-APR-2004.

24-SEP-2003; 2003WO-US030198.

27-SEP-2002; 2002US-0414227P.

18-NOV-2002; 2002US-0427594P.

07-FEB-2003; 2003US-0445724P.

07-MAR-2003; 2003US-0453277P.

(INCY-) INCYTE CORP.

Swarnakar A, Ho A, Hafalia AJA, Emerling BM, Ison CH, Nguyen DB;
 Chien D, Lu DAM, Yang YG, Yue H, Murage J, Griffin JA, Burrill JD;
 Wang JT, Marquis JP, Blake JJ, Chawla NK, Lee S, Becha SD, Lee SY;
 Richardson TW, Bhatia UG, Zheng W, Sprague WW, Tang YT, Azimzai Y;
 WPI; 2004-305155/28.
 N-PSDB; ADO44181.
 New human structural and cytoskeleton-associated proteins and
 polynucleotides for diagnosing, preventing or treating diseases
 associated with aberrant protein expression, e.g. cancer,
 atherosclerosis, HIV or stroke.

Claim 1; Page 150-152; 205pp; English.

The present sequence represents a structural and cytoskeleton-associated
 polypeptide (SCAP). The SCAP polypeptides and polynucleotides of the
 invention are useful for diagnosing, preventing or treating diseases or
 conditions associated with aberrant expression of SCAP, such as cell
 proliferative disorders (e.g. cancer or atherosclerosis), viral
 infections (e.g. HIV) or neurological disorders (e.g. Parkinson's
 disease, Alzheimer's disease or stroke). These may also be used for
 assessing the effects of exogenous compounds on the expression of nucleic
 acid and amino acid sequences of SCAP. The SCAP or its fragments are also
 useful in screening compounds for effectiveness as agonist or antagonist
 of the polypeptides, or in altering the expression of the target
 polynucleotide and compounds that specifically bind to or modulate the
 activity of the polypeptide.

Sequence 998 AA;

Query Match 31.6%; Score 559; DB 8; Length 998;

Best Local Similarity 38.9%; Pred. No. 7.1e-48;

Matches 132; Conservative 70; Mismatches 121; Indels 16; Gaps 7;

QY 19 RRVAVRLRPPF--VDGTAGAS-----DPPCYRGWDSCLSTANRNHQTLLK-YQDFAF 69

DB 11 QLMVALRVRPISVAELEGGATLIAHKVDQMVVLDPMDPDIDILRAHRSKSLFDVA 70

QY 70 YGERSTOODIYAGSVQPIRLHLLLEGONASVLAAYGTGAGKTHMLGSPQPGVPRALMD 129

DB 71 FDFATQEMVYQNTKSLIEGVISGNATVFAFGTGGCTYTLGLTQDQPGIYVQTLND 130


```

Qy 130 LQLTRREGARGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGNILIPGLSQX 189
Db 131 LFRAIETSD--MEYEVNSYLEIYNEMIRDLNPSLGLYELREDSKGVIVAGITEV 187
Qy 190 PISSFADPERHFLPASNRVTGATRLNQRSSRSRSHAVLLVKVDORERLAPRQ--REGKLY 247
Db 188 STINAKIMQLMGKNRQTEPAAQNTSSRSHAVLQVTVRQSRVKNILQEVROGRLF 247
Qy 248 LIDLAGSEDRNRRTKGLRLKESGAINTSLFVLGKVVVDAL--NOGLPR--VPYRDSKLTREL 305
Db 248 MIDLAGSERASQTRNQRMKEGAHINRSLALGNCINALSDKSNKYINVRDSKLTREL 307
Qy 306 QDSLGSAHLILANIPARFYLDTVSALNFAARSKEV 344
Db 308 KDSLGNSRTVMIAHISPASSAFESRNTLTLYAGRAKNI 346

RESULT 15
AAU76958
ID AAU76958 standard; protein; 357 AA.
XX AC
XX AAU76958;
XX DT
XX 21-MAY-2002 (first entry)
XX DE
XX Novel human kinesin motor protein, HsKip3d motor domain.
XX KW
XX Kinesin; motor protein; HsKip3d; cytosolic; immunosuppressive;
XX antiarthritic; antiinflammatory; vulnary; microtubule; inflammation;
XX cardiac hypertrophy; cellular proliferation disease; hyperplasia;
XX restenosis; cancer; solid tumour; carcinoma; autoimmune disease;
XX arthritis; graft rejection; inflammatory bowel disease; surgery;
XX angioplasty; nanotechnology; GeneChip array; expression monitoring;
XX human; enzyme.
XX OS
XX Homo sapiens.

Key Location/Qualifiers
FH 17..20
FT Region
FT /note= "Interacts with the adenine ring of ATP to form
FT part of the nucleotide-binding site"
FT 113..120
FT Region
FT /label= ATP/GTP binding_site_motif
FT 222..229
FT Region
FT /label= SwitchII motif
FT /note= "Responsible for sensing of the gamma phosphate
FT and initiating the conformational change between ATP and
FT ADP bound forms of the enzyme"
FT 258..263
FT Region
FT /label= SwitchIII motif
FT /note= "Responsible for sensing of the gamma phosphate
FT and initiating the conformational change between ATP and
FT ADP bound forms of the enzyme"
XX
XX W0200212268-A1.
XX
XX
XX 14-FEB-2002.
XX
XX 03-AUG-2001; 2001WO-US024285.
XX
XX 03-AUG-2000; 2000US-00632344.
XX
XX (CYTO-) CYTOKINETICS INC.
XX
XX Beraud C, Freedman R;
XX
XX WPI; 2002-217176/27.
XX
XX N-PSDB; ABK10348.
XX
XX Novel microtubule motor protein for screening modulators of HsKip3d,
XX useful in treatment of hyperproliferative disease e.g. cancer, autoimmune
XX disease, arthritis, graft rejection, inflammatory bowel disease.
XX

```

PS Claim 11; Fig 3; 66pp; English.

XX The invention describes an isolated novel human microtubule motor protein
 CC (1). (1) is useful for screening modulators of HsKip3d. The method
 CC comprises contacting (1) with a candidate agent in a test and control
 CC concentration, and assaying for the level of HsKip3d activity, where the
 CC HsKip3d activity is from binding activity or ATPase activity, and where a
 CC change in activity between the test and control concentration indicates a
 CC modulator. A composition containing (1) is useful in the treatment of e.g.
 CC inflammation, cardiac hypertrophy, cellular proliferation diseases e.g.
 CC hyperplasia, restenosis, cancer, including solid tumours of skin,
 CC breast, brain, lung, heart, bone, genitourinary tract, liver, nervous
 CC system, adrenal glands, haematologic, gynaecological and testicular
 CC carcinomas, autoimmune disease, arthritis, graft rejection, inflammatory
 CC bowel disease, proliferation induced after medical procedures, such as
 CC surgery, angioplasty. The motor domains may also be used in
 CC nanotechnological applications, and polynucleotides encoding the proteins
 CC is further useful for inclusion on GeneChip (RTM) array or for use in
 CC expression monitoring. This is the amino acid sequence of the novel human
 CC motor protein HsKip3d motor domain, described in the method of the
 CC invention

XX Sequence 357 AA;

Query Match 31.2%; Score 552.5; DB 5; Length 357;
 Best Local Similarity 39.7%; Pred. No. 6.9e-48;
 Matches 141; Conservative 59; Mismatches 114; Indels 41; Gaps 11;
 Qy 20 VRVAVRLRP-----FVDTAGASDP-----PCVRGMDSCSLEIANWRNHQ 59
 Db 12 MKYVVRVRPENTKEAAGFKVHVVDKHLVDPKQEEVSFFHGKKTNNQNVIKQN-- 69
 Qy 60 ETLKYOFDAFYGERSTQDDIYAGSVOPILRHLLLEGQNASVLYAGPTGAGKTHMLGSPQ 119
 Db 70 KDLKFVDAVDFDTSTQSEVFHTTRPILRSFLNGYNTVLTATGATGAGKTHMLGSADE 129
 Qy 120 PGVIPRALMDLLQLTREGAEGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRG 179
 Db 130 PGVYMLTMLHLKCMDEIKEEK--ICSTAVSYLEVYNEQIRDLL--VNSGPLAVREDTQK 185
 Qy 180 NILIPGLS-QKPISSFADPERHFLP-ASRNVTVGATRLNQRSSRSRSHAVLLVKVDORERLA 237
 Db 186 GVVVHGLTLHQPKSS--EELHLLDNGKNRKTQPTDMNATSSRSRSHAVFOILRQODKTA 243
 Qy 238 PFRQ--REGKLYLIDLAGSEDRNRRTKGLRLKESGAINTSLFVLGKVVVDAL-----NO 289
 Db 244 SINQNVRIAKMSLIDLAGSERASTSGAKGTRFVEGTNINRSLALGNVINIALADSKRKQ 303
 Qy 290 GLPRVPYRDSKLTRELQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARSKEV 344
 Db 304 ---HIPYRNSKLTRELKDSLGGNCQITMIAAVSPSSVFYDDTYNTLKYANRAKDI 355

Search completed: November 5, 2004, 18:44:40
 Job time : 72.0157 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 04:57:04 ; Search time 94.4529 Seconds
(without alignments)
7833.865 Million cell updates/sec

Title: US-10-797-893-5
Perfect score: 1041
Sequence: 1 atgggtcgctgcggtaag.....ccaggaggatgaattga 1041

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A-COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B-COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A-COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B-COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS-COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1041	100.0	1041	3	US-09-724-224-5
2	1041	100.0	1041	4	US-10-093-317-5
3	1038.4	99.8	1464	3	US-09-724-224-7
4	1038.4	99.8	1464	4	US-10-093-317-7
5	1038	99.7	1115	3	US-09-724-224-1
6	1038	99.7	1115	4	US-10-093-317-1
7	1035.4	99.5	1538	3	US-09-724-224-3
8	1035.4	99.5	1538	4	US-10-093-317-3
9	1035.4	99.5	2097	4	US-09-595-684B-34
10	136.6	13.1	1152	4	US-09-883-096-6
11	136.6	13.1	4108	4	US-09-883-096-1
12	127.8	12.3	897	3	US-09-621-233-1
13	127.8	12.3	897	3	US-09-724-508-1
14	127.8	12.3	897	3	US-09-724-516-1
15	127.8	12.3	897	3	US-10-090-695-1
16	125.2	12.0	1026	3	US-09-641-806-1
17	125.2	12.0	1026	4	US-09-723-129-1
18	125.2	12.0	1026	4	US-09-722-862-1
19	124.2	11.9	1011	3	US-09-641-806-3
20	124.2	11.9	1011	4	US-09-723-129-3
21	124.2	11.9	1011	4	US-09-722-862-3
22	121.4	11.7	1014	4	US-09-883-096-3
23	98.6	9.5	3572	2	US-08-713-815A-2
24	93.8	9.0	3258	4	US-09-799-451-586
25	85.6	8.2	4014	3	US-09-541-782-1
26	85.6	8.2	4014	4	US-09-723-820-1
27	85.6	8.2	4014	4	US-10-270-085-1

28	79.2	7.6	1032	4	US-09-967-908A-7	Sequence 7, Appli
29	79.2	7.6	1032	4	US-10-159-151-7	Sequence 7, Appli
30	79.2	7.6	1092	4	US-09-967-908A-9	Sequence 9, Appli
31	79.2	7.6	1092	4	US-10-159-151-9	Sequence 9, Appli
32	79.2	7.6	1122	4	US-09-967-908A-3	Sequence 3, Appli
33	79.2	7.6	1122	4	US-10-159-151-3	Sequence 3, Appli
34	79.2	7.6	1146	4	US-09-967-908A-5	Sequence 5, Appli
35	79.2	7.6	1146	4	US-10-159-151-5	Sequence 5, Appli
36	79.2	7.6	6409	4	US-09-967-908A-1	Sequence 1, Appli
37	79.2	7.6	6409	4	US-10-159-151-1	Sequence 1, Appli
38	78.4	7.5	1065	3	US-09-724-511-3	Sequence 3, Appli
39	78.4	7.5	1065	4	US-09-723-097-3	Sequence 3, Appli
40	78.4	7.5	1065	4	US-09-632-344-3	Sequence 3, Appli
41	78.4	7.5	1217	3	US-09-724-511-1	Sequence 1, Appli
42	78.4	7.5	1217	4	US-09-723-097-1	Sequence 1, Appli
43	78.4	7.5	1217	4	US-09-632-344-1	Sequence 1, Appli
44	77.8	7.5	3930	3	US-09-162-373-2	Sequence 2, Appli
45	77.8	7.5	3930	3	US-09-467-946-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-724-224-5
; Sequence 5, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09/724, 224
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597, 292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Human
US-09-724-224-5

Query Match	100.0%	Score 1041;	DB 3;	Length 1041;
Best Local Similarity	100.0%	Pred. No. 0;	Mismatches 0;	Indels 0; Gaps 0;
Matches 1041;	Conservative	0;		
QY	1	ATGGTTCGCTGTGGCTAAGCAAGATTGGAGTACTCGTCCACCTCCAGCTCGCGTA	60	
Db	1	ATGGTTCGCTGTGGCTAAGCAAGATTGGAGTACTCGTCCACCTCCAGCTCGCGTA	60	
QY	61	AGGTGGCTGTGGCTGTGGCTAAGCAAGATTGGAGTACTCGTCCACCTCCAGCTCGCGTA	120	
Db	61	AGGTGGCTGTGGCTGTGGCTAAGCAAGATTGGAGTACTCGTCCACCTCCAGCTCGCGTA	120	
QY	121	TGTGTGGCTGTGGCTGTGGCTAAGCAAGATTGGAGTACTCGTCCACCTCCAGCTCGCGTA	180	
Db	121	TGTGTGGCTGTGGCTGTGGCTAAGCAAGATTGGAGTACTCGTCCACCTCCAGCTCGCGTA	180	
QY	181	ACTCTCAATACCAAGTTTGTATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTAT	240	
Db	181	ACTCTCAATACCAAGTTTGTATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTAT	240	
QY	241	GCAGTTTCAGTGTGGCTGTGGCTAAGCAAGATTGGAGTACTCGTCCACCTCCAGCTCGCGTA	300	
Db	241	GCAGTTTCAGTGTGGCTGTGGCTAAGCAAGATTGGAGTACTCGTCCACCTCCAGCTCGCGTA	300	
QY	301	GCCTATGAGCCACAGGAGTGTGGAGACGACACCAATGCTGGGAGAGAGGAGCAACT	360	
Db	301	GCCTATGAGCCACAGGAGTGTGGAGACGACACCAATGCTGGGAGAGAGGAGCAACT	360	
QY	361	GGGTGTATCCCGGGGCTCTCATGACCTCTGACGCTCACAAAGGAGGAGGAGGCTCGCGAG	420	

```

Db 361 GGGGTGATCCGGGGCTCTCATGGACCTCTGAGCTCAAGGGAGGGGTGCCGAG 420
Qy 421 GGGGGCCATGGGCCCTTTCTGTACCATGCTCTTACCTAGAGATCTACAGGAGAGGTA 480
Db 421 GGGGGCCATGGGCCCTTTCTGTACCATGCTCTTACCTAGAGATCTACAGGAGAGGTA 480
Qy 481 TTAGACTCTGAGACCTGCTTGGGAGACCTGGTAATCCGAGAGACTGCCGGGGAAT 540
Db 481 TTAGACTCTGAGACCTGCTTGGGAGACCTGGTAATCCGAGAGACTGCCGGGGAAT 540
Qy 541 ATCTGATTCGGGGTCTCTCCAGAGACCATCAGTAGCTTTGCTGATTTGAGGGGCA 600
Db 541 ATCTGATTCGGGGTCTCTCCAGAGACCATCAGTAGCTTTGCTGATTTGAGGGGCA 600
Qy 601 TTCTGCGAGCCAGTCCGAATCGGACTGTAGGAGCCACCCGGCTCAACAGGCTCTCTCC 660
Db 601 TTCTGCGAGCCAGTCCGAATCGGACTGTAGGAGCCACCCGGCTCAACAGGCTCTCTCC 660
Qy 661 CGCAGTCATGCTGCTGCTCTGCTCAAGGTGGACCGAGGGGAAAGCTTTGGCCCAATTCGC 720
Db 661 CGCAGTCATGCTGCTGCTCTGCTCAAGGTGGACCGAGGGGAAAGCTTTGGCCCAATTCGC 720
Qy 721 CAGCGAGAGGGAATCTTACCTGATTGACTTTGGCTGGGTCAAGGAGCAACCGCGGCA 780
Db 721 CAGCGAGAGGGAATCTTACCTGATTGACTTTGGCTGGGTCAAGGAGCAACCGCGGCA 780
Qy 781 GGCACAAAGGGCTTCGGCTAAAGAGAGTGAGGAGCAACACCTCCCTGTTGCTG 840
Db 781 GGCACAAAGGGCTTCGGCTAAAGAGAGTGAGGAGCAACACCTCCCTGTTGCTG 840
Qy 841 GGCACAAAGGGCTTCGGCTAAAGAGAGTGAGGAGCAACACCTCCCTGTTGCTG 900
Db 841 GGCACAAAGGGCTTCGGCTAAAGAGAGTGAGGAGCAACACCTCCCTGTTGCTG 900
Qy 901 CTCACCTGCTTATGAGAGCTCTCTGGGTGGCTCAGCCCAAGTATCTTATGCCAAC 960
Db 901 CTCACCTGCTTATGAGAGCTCTCTGGGTGGCTCAGCCCAAGTATCTTATGCCAAC 960
Qy 961 ATTGGCCCTGAGAGAGCTTCTTACCTAGACAGCTCCGGACTCAACTTTGCTGCCAGG 1020
Db 961 ATTGGCCCTGAGAGAGCTTCTTACCTAGACAGCTCCGGACTCAACTTTGCTGCCAGG 1020
Qy 1021 TCCAGGAGGTGATCAATTGA 1041
Db 1021 TCCAGGAGGTGATCAATTGA 1041

```

RESULT 2

```

US-10-093-317-5
; Sequence 5, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093,317
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Human
US-10-093-317-5

```

```

Query Match 100.0%; Score 1041; DB 4; Length 1041;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 ATGGGTGCTGCTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCCGCTA 60
Db 1 ATGGGTGCTGCTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCCGCTA 60
Qy 61 AGGGTGGCTGTGCGACTGCGGCCATTTGTGGATGGAACAGCGGGAGCAAGTGTATCCCCC 120
Db 61 AGGGTGGCTGTGCGACTGCGGCCATTTGTGGATGGAACAGCGGGAGCAAGTGTATCCCCC 120
Qy 121 TGTGTGCGGGGATGAGCAGCTGCTCTTAGAGATGCTTAAGTGGAGAACCCAGGAG 180
Db 121 TGTGTGCGGGGATGAGCAGCTGCTCTTAGAGATGCTTAAGTGGAGAACCCAGGAG 180
Qy 181 ACTCTCAAAATACCAAGTTTGTATGTCCTTCTATGGGAGAGAGTACTCAGCAGGACATCTAT 240
Db 181 ACTCTCAAAATACCAAGTTTGTATGTCCTTCTATGGGAGAGAGTACTCAGCAGGACATCTAT 240
Qy 241 GCAGTTTCAAGTGGAGCCCATCTTAAGGCACTTGTGGAGGGGCAAAATGCGAGTGTGTT 300
Db 241 GCAGTTTCAAGTGGAGCCCATCTTAAGGCACTTGTGGAGGGGCAAAATGCGAGTGTGTT 300
Qy 301 GCCTATGAGCCACAGAGAGCTGGGAGACGACACAAATGCTGGGAGGAGGAGGAGCACT 360
Db 301 GCCTATGAGCCACAGAGAGCTGGGAGAGCGACAAATGCTGGGAGGAGGAGGAGCACT 360
Qy 361 GGGGTGATCCCGGGGCTCTCATGGAATCTCTGAGCTCAAGGGAGGAGGGTGGCGAG 420
Db 361 GGGGTGATCCCGGGGCTCTCATGGAATCTCTGAGCTCAAGGGAGGAGGGTGGCGAG 420
Qy 421 GGGGGCCATGGGGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTA 480
Db 421 GGGGGCCATGGGGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTA 480
Qy 481 TTAGACTCTCTGGAACCTGCTTGGGAGAGCTGGTAATCCGAGAGAGTCTCGGGGGAAT 540
Db 481 TTAGACTCTCTGGAACCTGCTTGGGAGAGCTGGTAATCCGAGAGAGTCTCGGGGGAAT 540
Qy 541 ATCTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGCGCAC 600
Db 541 ATCTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGCGCAC 600
Qy 601 TTCTGCGAGCCAGTCCGAATCGGACTGTAGGAGCCACCCGGCTCAACAGGCTCTCTCC 660
Db 601 TTCTGCGAGCCAGTCCGAATCGGACTGTAGGAGCCACCCGGCTCAACAGGCTCTCTCC 660
Qy 661 CGCAGTCACTGCTGCTCTGCTGTCAGGAGGAGCAAGGAGGAGGAGGAGGAGGAGGAG 720
Db 661 CGCAGTCACTGCTGCTCTGCTGTCAGGAGGAGCAAGGAGGAGGAGGAGGAGGAGGAG 720
Qy 721 CAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
Db 721 CAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
Qy 781 GGCACAAAGGGCTTCGGCTAAAGAGAGTGAGGAGCCATCAACACCTCTCTTGTGCTG 840
Db 781 GGCACAAAGGGCTTCGGCTAAAGAGAGTGAGGAGCCATCAACACCTCTCTTGTGCTG 840
Qy 841 GGCACAAAGGGCTTCGGCTAAAGAGAGTGAGGAGCCATCAACACCTCTCTTGTGCTG 900
Db 841 GGCACAAAGGGCTTCGGCTAAAGAGAGTGAGGAGCCATCAACACCTCTCTTGTGCTG 900
Qy 901 CTCACCTGCTTATGAGAGCTCTCTGGGTGGCTCAGCCCAAGTATCTTATGCCAAC 960
Db 901 CTCACCTGCTTATGAGAGCTCTCTGGGTGGCTCAGCCCAAGTATCTTATGCCAAC 960
Qy 961 ATTGGCCCTGAGAGAGCTTCTTACCTAGACAGCTCCGGACTCAACTTTGCTGCCAGG 1020
Db 961 ATTGGCCCTGAGAGAGCTTCTTACCTAGACAGCTCCGGACTCAACTTTGCTGCCAGG 1020
Qy 1021 TCCAGGAGGTGATCAATTGA 1041
Db 1021 TCCAGGAGGTGATCAATTGA 1041

```

```

RESULT 3
US-09-724-224-7
; Sequence 7, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Human
US-09-724-224-7

Query Match          99.8%; Score 1038.4; DB 3; Length 1464;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTA 60
DB 1 ATGGGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTA 60
QY 61 AGGTGCTGTCGGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTA 120
DB 61 AGGTGCTGTCGGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTA 120
QY 121 TGTGTCGGGGATGACAGCTGCTCTTAGAGATTGCTTAACTGGAGAACCAACAGGAG 180
DB 121 TGTGTCGGGGATGACAGCTGCTCTTAGAGATTGCTTAACTGGAGAACCAACAGGAG 180
QY 181 ACTCTCAATATACAGTTTGTATGCTTCTATGGGGAGAGGAGTACTCAGCAGGACATCTAT 240
DB 181 ACTCTCAATATACAGTTTGTATGCTTCTATGGGGAGAGGAGTACTCAGCAGGACATCTAT 240
QY 241 GCAGGTTCAAGTGGCCCATCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTA 300
DB 241 GCAGGTTCAAGTGGCCCATCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTA 300
QY 421 GCGCGGCCATGCGGCCCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTA 480
DB 421 GCGCGGCCATGCGGCCCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTA 480
QY 481 TTAGACCTCTGACCCCTGCTTGGGAGAGCTGTTAACTCCGAGAGAGCTGCGCGGGGAT 540
DB 481 TTAGACCTCTGACCCCTGCTTGGGAGAGCTGTTAACTCCGAGAGAGCTGCGCGGGGAT 540
QY 541 ATCTCTGATTCCGGGCTCTCCAGAGGCCCATCAGTAGCTTTGCTGATTTTTCAGGGGAC 600
DB 541 ATCTCTGATTCCGGGCTCTCCAGAGGCCCATCAGTAGCTTTGCTGATTTTTCAGGGGAC 600
QY 601 TTCCTGCCAGCAGTCGAATCGGACTGTAGAGGACCCCGGCTCAACAGCGCTCTTCC 660
DB 601 TTCCTGCCAGCAGTCGAATCGGACTGTAGAGGACCCCGGCTCAACAGCGCTCTTCC 660
QY 661 CCGAGTCATGCTGCTCTGCTCAAGGTGGACCGGAGGAGCTTTGGCCCCCATTTGCG 720
DB 661 CCGAGTCATGCTGCTCTGCTCAAGGTGGACCGGAGGAGCTTTGGCCCCCATTTGCG 720

```

```

QY 721 CAGCGAGAGGGAAAACTCTACCTGATTGATTGGCTGGGTGAGAGCAACCGCGCGACA 780
DB 721 CAGCGAGAGGGAAAACTCTACCTGATTGATTGGCTGGGTGAGAGCAACCGCGCGACA 780
QY 781 GCGAACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTG 840
DB 781 GCGAACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTG 840
QY 841 GCGAAAGTGTAGTGGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGGACAGCAAG 900
DB 841 GCGAAAGTGTAGTGGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGGACAGCAAG 900
QY 901 CTCACCTCGCTATTGCGAGACTCTCTGCGGTGGCTCAGCCACAGTATCTTATTGCGAAC 960
DB 901 CTCACCTCGCTATTGCGAGACTCTCTGCGGTGGCTCAGCCACAGTATCTTATTGCGAAC 960
QY 961 ATTGCGCCCTGAGAGAGGCTTCTACCTAGACACAGTCTCGGACTCAACTTTGCTGCGAGG 1020
DB 961 ATTGCGCCCTGAGAGAGGCTTCTACCTAGACACAGTCTCGGACTCAACTTTGCTGCGAGG 1020
QY 1021 TCCAAGGAGGTGATCAATTG 1040
DB 1021 TCCAAGGAGGTGATCAATCG 1040

RESULT 4
US-10-093-317-7
; Sequence 7, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093,317
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Human
US-10-093-317-7

Query Match          99.8%; Score 1038.4; DB 4; Length 1464;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTA 60
DB 1 ATGGGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTA 60
QY 61 AGGTGCTGTCGGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTA 120
DB 61 AGGTGCTGTCGGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTA 120
QY 121 TGTGTCGGGGATGACAGCTGCTCTTAGAGATTGCTTAACTGGAGAACCAACAGGAG 180
DB 121 TGTGTCGGGGATGACAGCTGCTCTTAGAGATTGCTTAACTGGAGAACCAACAGGAG 180
QY 181 ACTCTCAATATACAGTTTGTATGCTTCTATGGGGAGAGGAGTACTCAGCAGGACATCTAT 240
DB 181 ACTCTCAATATACAGTTTGTATGCTTCTATGGGGAGAGGAGTACTCAGCAGGACATCTAT 240
QY 241 GCAGGTTCAAGTGGCCCATCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTA 300
DB 241 GCAGGTTCAAGTGGCCCATCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTA 300
QY 301 GCCTATGAGACCCAGGAGCTGGGAGAGCGCACAAATGCTGGGAGGCCAGGCAACCT 360
DB 301 GCCTATGAGACCCAGGAGCTGGGAGAGCGCACAAATGCTGGGAGGCCAGGCAACCT 360
QY 361 GGGGTGATCCCGGGGCTCTCATGACCTCTCTGACAGTCAACAGGGAGGAGGTCGCGAG 420
DB 361 GGGGTGATCCCGGGGCTCTCATGACCTCTCTGACAGTCAACAGGGAGGAGGTCGCGAG 420
QY 421 GCGCGGCCATGCGGCCCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTA 480
DB 421 GCGCGGCCATGCGGCCCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTA 480
QY 481 TTAGACCTCTGACCCCTGCTTGGGAGAGCTGTTAACTCCGAGAGAGCTGCGCGGGGAT 540
DB 481 TTAGACCTCTGACCCCTGCTTGGGAGAGCTGTTAACTCCGAGAGAGCTGCGCGGGGAT 540
QY 541 ATCTCTGATTCCGGGCTCTCCAGAGGCCCATCAGTAGCTTTGCTGATTTTTCAGGGGAC 600
DB 541 ATCTCTGATTCCGGGCTCTCCAGAGGCCCATCAGTAGCTTTGCTGATTTTTCAGGGGAC 600
QY 601 TTCCTGCCAGCAGTCGAATCGGACTGTAGAGGACCCCGGCTCAACAGCGCTCTTCC 660
DB 601 TTCCTGCCAGCAGTCGAATCGGACTGTAGAGGACCCCGGCTCAACAGCGCTCTTCC 660
QY 661 CCGAGTCATGCTGCTCTGCTCAAGGTGGACCGGAGGAGCTTTGGCCCCCATTTGCG 720
DB 661 CCGAGTCATGCTGCTCTGCTCAAGGTGGACCGGAGGAGCTTTGGCCCCCATTTGCG 720

```

QY 361 GGGGTGATCCCGGGCTCTCATGGACCTCTGACCTCTGACCTCACAAAGGAGGAGGTGCGGAG 420
 Db 361 GGGGTGATCCCGGGCTCTCATGGACCTCTGACCTCTGACCTCACAAAGGAGGAGGTGCGGAG 420
 QY 421 GGGCGGCATGCGGCTCTCTCTGTCACCATGTCTTACCTAGAGATCTACCAAGAGAGGTA 480
 Db 421 GGGCGGCATGCGGCTCTCTCTGTCACCATGTCTTACCTAGAGATCTACCAAGAGAGGTA 480
 QY 481 TTAGACCTCTGGAACCTGCTCTGCGGAGACCTGGAATCCGAGAAGACTGCGCGGGGAAT 540
 Db 481 TTAGACCTCTGGAACCTGCTCTGCGGAGACCTGGAATCCGAGAAGACTGCGCGGGGAAT 540
 QY 541 ATCTGATTTCCGGGTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTTTCAGCGGCAC 600
 Db 541 ATCTGATTTCCGGGTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTTTCAGCGGCAC 600
 QY 601 TTCTGCCAGCCAGTCGAATCGGACTGTAGAGCCACCCGGCTCAACAGCGCTCTCTCC 660
 Db 601 TTCTGCCAGCCAGTCGAATCGGACTGTAGAGCCACCCGGCTCAACAGCGCTCTCTCC 660
 QY 661 CGCAGTCATGCTGCTGCTCTGCTCAAGGTGAGCCAGCGGGAACGTTTGGCCCATTTCCG 720
 Db 661 CGCAGTCATGCTGCTGCTCTGCTCAAGGTGAGCCAGCGGGAACGTTTGGCCCATTTCCG 720
 QY 721 CAGCGAGAGGGAATCTCTACCTGATTCAGTTGGCTGGGTGAGGACCAACCGGCGCAC 780
 Db 721 CAGCGAGAGGGAATCTCTACCTGATTCAGTTGGCTGGGTGAGGACCAACCGGCGCAC 780
 QY 781 GSCAAAGGGCTTTCCGCTAAAGAGAGTGAGGACCATCAACACCTCTGTTTGTCTG 840
 Db 781 GSCAAAGGGCTTTCCGCTAAAGAGAGTGAGGACCATCAACACCTCTGTTTGTCTG 840
 QY 841 GSCAAAGTGAGTGGCTGATCAGGCTTCAAGGCTTCCCTGCTGATCTTATTCGGGACAGGA 900
 Db 841 GSCAAAGTGAGTGGCTGATCAGGCTTCAAGGCTTCCCTGCTGATCTTATTCGGGACAGGA 900
 QY 901 CTCACCTGCGCTATTGAGGACTCTCTGGTGGCTCAGCCACAGTATCTTATTCGCAAC 960
 Db 901 CTCACCTGCGCTATTGAGGACTCTCTGGTGGCTCAGCCACAGTATCTTATTCGCAAC 960
 QY 961 ATTGCCCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGG 1020
 Db 961 ATTGCCCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGG 1020
 QY 1021 TCCAGGAGGTGATCAATTG 1040
 Db 1021 TCCAGGAGGTGATCAATCG 1040

RESULT 5

US-09-724-224-1
 ; Sequence 1, Application US/09724224
 ; Patent No. 6387644
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
 ; TITLE OF INVENTION: their use
 ; FILE REFERENCE: 1044
 ; CURRENT APPLICATION NUMBER: US/09/724,224
 ; CURRENT FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: 09/597,292
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1115
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-724-224-1

Query Match 99.7%; Score 1038; DB 3; Length 1115;
 Best Local Similarity 100.0%; Pred. No. 0;

Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 GGTGCTCTCGGCTAAGCAAGATTGGAGTACTCGTCTGCTCACTCTCCAGCTGCGTAAGG 63
 Db 78 GGTGCTCTCGGCTAAGCAAGATTGGAGTACTCGTCTGCTCACTCTCCAGCTGCGTAAGG 137
 QY 64 GTGGCTGTGCACACTGCGGCCATTTGTGATGGAAACAGCGGGGAGCAAGTGTATCCCTTGT 123
 Db 138 GTGGCTGTGCACACTGCGGCCATTTGTGATGGAAACAGCGGGGAGCAAGTGTATCCCTTGT 197
 QY 124 GTGGGGGCATGAGCAGCTGCTCTCTAGAGATTGCTTAACCTGAGGAACCAACAGAGACT 183
 Db 198 GTGGGGGCATGAGCAGCTGCTCTCTAGAGATTGCTTAACCTGAGGAACCAACAGAGACT 257
 QY 184 CTCAAAATACAGTTTGTATGCTTCTATGGGAGAGGAGTACTCTAGCAGGACATCTATGCA 243
 Db 258 CTCAAAATACAGTTTGTATGCTTCTATGGGAGAGGAGTACTCTAGCAGGACATCTATGCA 317
 QY 244 GGTTCAGTGCAGCCATCTTAAGGCACCTTCTGAGGAGGAGGAGTATGCGAGTGTCTTGGC 303
 Db 318 GGTTCAGTGCAGCCATCTTAAGGCACCTTCTGAGGAGGAGGAGTATGCGAGTGTCTTGGC 377
 QY 304 TATGGAACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGGAGGAGGAGGAGGAGG 363
 Db 378 TATGGAACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGGAGGAGGAGGAGGAGG 437
 QY 364 GTGATCCCGCGGCTCTCATGGAACCTCTGAGCTCAACAGGAGGAGGAGGAGGAGGAGG 423
 Db 438 GTGATCCCGCGGCTCTCATGGAACCTCTGAGCTCAACAGGAGGAGGAGGAGGAGGAGG 497
 QY 424 CGGCCATGGGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGGAGGAGGAGGAGG 483
 Db 498 CGGCCATGGGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGGAGGAGGAGGAGG 557
 QY 484 GACCTCTGGAACCTGCTCTGGGAGACTGCTGTAATCCGAGAGAGTCTGCGGGGAGGAGTATC 543
 Db 558 GACCTCTGGAACCTGCTCTGGGAGACTGCTGTAATCCGAGAGAGTCTGCGGGGAGGAGTATC 617
 QY 544 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTTTCAGGCGGAGTCTTC 603
 Db 618 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTTTCAGGCGGAGTCTTC 677
 QY 604 CTGCGAGCCAGTGGAAATCGGAGCTGAGGAGCCACCGGCTCAACAGGAGTCTCTCCGCG 663
 Db 678 CTGCGAGCCAGTGGAAATCGGAGCTGAGGAGCCACCGGCTCAACAGGAGTCTCTCCGCG 737
 QY 664 AGTCATGCTGTGCTCTGCTCAAGGTGGAACAGCGGAGGAGGAGGAGGAGGAGGAGGAG 723
 Db 738 AGTCATGCTGTGCTCTGCTCAAGGTGGAACAGCGGAGGAGGAGGAGGAGGAGGAGGAG 797
 QY 724 CGAGAGGGAACCTCTACCTGATTTGACTTGGTGGGTTCAGAGGACCAACCGGCGCACAGGC 783
 Db 798 CGAGAGGGAACCTCTACCTGATTTGACTTGGTGGGTTCAGAGGACCAACCGGCGCACAGGC 857
 QY 784 AACAGAGGCTTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTTGTCTTGGGC 843
 Db 858 AACAGAGGCTTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTTGTCTTGGGC 917
 QY 844 AAAGTGTAGATGCTGGAATTCAGGCGCTCTCTGCTGTAACCTATCGGAGCAGCAAGCTC 903
 Db 918 AAAGTGTAGATGCTGGAATTCAGGCGCTCTCTGCTGTAACCTATTCGGAACAGCAAGCTC 977
 QY 904 ACTCGCTATTGAGGAGTCTCTGGGTGGGTTCAGGCGCCACAGTATCTTATTCGCAACAT 963
 Db 978 ACTCGCTATTGAGGAGTCTCTGGGTGGGTTCAGGCGCCACAGTATCTTATTCGCAACAT 1037
 QY 964 GCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTTGTGCGAGGTCC 1023
 Db 1038 GCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTTGTGCGAGGTCC 1097
 QY 1024 AAGAGGTGATCAATTGA 1041
 Db 1098 AAGAGGTGATCAATTGA 1115

Qy	724	CGAGAGGAAAACTCTACCTGATGTACTTGGCTGGGTACAGAGGACAAACCGGCGCACAGGC	783
Db	798	CGAGAGGAAAACTCTACCTGATGTACTTGGCTGGGTACAGAGGACAAACCGGCGCACAGGC	857
Qy	784	AACAAGGGCCCTCGGCTAAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCTGGCC	843
Db	858	AACAAGGGCCCTCGGCTAAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCTGGCC	917
Qy	844	AAAGTGGTAGATGCGCTGAATCAGGGCCTCCCTCGTGTACTTATCGGACAGCAAGCTC	903
Db	918	AAAGTGGTAGATGCGCTGAATCAGGGCCTCCCTCGTGTACTTATCGGACAGCAAGCTC	977
Qy	904	ACTCGCCTATTGACGAGACTCTCTGGTGGCTCAGCCACACAGTATCCTTATTTGCCAACATT	963
Db	978	ACTCGCCTATTGACGAGACTCTCTGGTGGCTCAGCCACACAGTATCCTTATTTGCCAACATT	1037
Qy	964	GCCCTGAGAGACGCTTCTACCTAGACACAGTCTCCGACCTCAACTTTTGCTGCCAGTCC	1023
Db	1038	GCCCTGAGAGACGCTTCTACCTAGACACAGTCTCCGACCTCAACTTTTGCTGCCAGTCC	1097
Qy	1024	AAGGAGGTGATCAATTGA	1041
Db	1098	AAGGAGGTGATCAATTGA	1115
RESULT 7			
US-09-724-224-3			
; Sequence 3, Application US/09724224			
; Patent No. 6387644			
; GENERAL INFORMATION:			
; APPLICANT: Beraud, Christophe			
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for			
; TITLE OF INVENTION: their use			
; FILE REFERENCE: 1044			
; CURRENT APPLICATION NUMBER: US/09/724,224			
; CURRENT FILING DATE: 2000-11-28			
; PRIOR APPLICATION NUMBER: 09/597,292			
; PRIOR FILING DATE: 2000-06-20			
; NUMBER OF SEQ ID NOS: 8			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 3			
; LENGTH: 1538			
; TYPE: DNA			
; ORGANISM: Human			
US-09-724-224-3			
Query Match 99.5%; Score 1035.4; DB 3; Length 1538;			
Best Local Similarity 99.9%; Pred. No. 9.9e-315;			
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	4	GGTGCTGTGCGCTAAGCAAGATTGGAGTACTCGTGTCCACCTCCAGCTCGGTTAAGG	63
Db	78	GGTGTGCTGTGCGCTAAGCAAGATTGGAGTACTCGTGTCCACCTCCAGCTCGGTTAAGG	137
Qy	64	GTGCGCTGTGCGACTCGCGCCATTGTGTGATGGAAACAGCGGAGCAAGTGATCCCCCTGT	123
Db	138	GTGCGCTGTGCGACTCGCGCCATTGTGTGATGGAAACAGCGGAGCAAGTGATCCCCCTGT	197
Qy	124	GTGCGGGCATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACT	183
Db	198	GTGCGGGCATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACT	257
Qy	184	CTCAATACAGTTTGATGCTTCTATGGGGAGAGAGTACTCAGCAAGGACATCTATGCA	243
Db	258	CTCAATACAGTTTGATGCTTCTATGGGGAGAGAGTACTCAGCAAGGACATCTATGCA	317
Qy	244	GGTTTCAGTCAGCCCATCTTAAGGCACTTGCTCGAAGGGCAGATGCTCAGTGTGCTTGCC	303
Db	318	GGTTTCAGTCAGCCCATCTTAAGGCACTTGCTCGAAGGGCAGATGCTCAGTGTGCTTGCC	377
Qy	304	TATGGACCCACAGGAGCTGGGAAGAGCAACAATGCTGGGCGAGCCACAGACCACTGGG	363


```

Db 378 TATGACCCACAGGAGCTGGGAAGAGCGACACAAATGCTGGGCGAGCCGAGAGCAACTGGG 437
QY 364 GTGATCCCGCGGCTCTCATGACCTCTGAGCTTCAAGAGGAGAGGTTGCCAGGCG 423
Db 438 GTGATCCCGCGGCTCTCATGACCTCTGAGCTTCAAGAGGAGAGGTTGCCAGGCG 497
QY 424 CGGCCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTATTA 483
Db 498 CGGCCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTATTA 557
QY 484 GACCTCTGAGCCCTGCTCGGAGACCTGGTAATCCGAGAGAGCTGCGGGGGAAATATC 543
Db 558 GACCTCTGAGCCCTGCTCGGAGACCTGGTAATCCGAGAGAGCTGCGGGGGAAATATC 617
QY 544 CTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTTCAGCGGCACTTC 603
Db 618 CTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTTCAGCGGCACTTC 677
QY 604 CTGCCAGCCAGTCGAATCGGACTGTAGGAGCCACCCGCTCAACAGGGCTCTCCGCG 663
Db 678 CTGCCAGCCAGTCGAATCGGACTGTAGGAGCCACCCGCTCAACAGGGCTCTCCGCG 737
QY 664 AGTCATGCTGCTCTCTGCTCAAGGTGGACAGCGGGAACGTTTGGCCGCCATTTCCGCG 723
Db 738 AGTCATGCTGCTCTCTGCTCAAGGTGGACAGCGGGAACGTTTGGCCGCCATTTCCGCG 797
QY 724 CGAGAGGGAAATCTTACCTGATGACTTTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 783
Db 798 CGAGAGGGAAATCTTACCTGATGACTTTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 857
QY 784 AACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCTGTTGCTGCTGGG 843
Db 858 AACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCTGTTGCTGCTGGG 917
QY 844 AAAGTGTAGTGTGCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTTCAGCGGCACTTC 903
Db 918 AAAGTGTAGTGTGCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTTCAGCGGCACTTC 977
QY 904 ACTGCGCTATTGACAGACTCTCTGGGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAG 963
Db 978 ACTGCGCTATTGACAGACTCTCTGGGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAG 1037
QY 964 GCGCCCTGAGAGCGCTTCTACCTAGACACAGTCTCGGCACTCAACTTGTGCTGCGGCTCC 1023
Db 1038 GCGCCCTGAGAGCGCTTCTACCTAGACACAGTCTCGGCACTCAACTTGTGCTGCGGCTCC 1097
QY 1024 AAGGAGGTGATCAATTG 1040
Db 1098 AAGGAGGTGATCAATCG 1114

```

RESULT 8

```

US-10-093-317-3
; Sequence 3, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093,317
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: Human
US-10-093-317-3

```

```

Query Match 99.5%; Score 1035.4; DB 4; Length 1538;
Best Local Similarity 99.9%; Pred. No. 9.9e-315;
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 GGTGCTCTCGGCTAAAGCAAGATTGGAGCTACTGTCGCTCCACCTCCAGCTCGCTAAGG 63
Db 78 GGTGCTCTCGGCTAAAGCAAGATTGGAGCTACTGTCGCTCCACCTCCAGCTCGCTAAGG 137
QY 64 GTGGCTGTGCACTGCGGCCCATTTGTGATGGAACAGCGGAGCAAGTGTATCCCTCTGT 123
Db 138 GTGGCTGTGCACTGCGGCCCATTTGTGATGGAACAGCGGAGCAAGTGTATCCCTCTGT 197
QY 124 GTGGGCGCATGGAAGCTGCTCTAGAGATTGCTAATCTGAGGAAACACAGGAGACT 183
Db 198 GTGGGCGCATGGAAGCTGCTCTAGAGATTGCTAATCTGAGGAAACACAGGAGACT 257
QY 184 CTCAATAACCAAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 243
Db 258 CTCAATAACCAAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 317
QY 244 GGTTCAGTGTGAGCCCATCTTAAGGCACTTGTGTAAGGGGCAAGTGTGCTGCTGCTGCC 303
Db 318 GGTTCAGTGTGAGCCCATCTTAAGGCACTTGTGTAAGGGGCAAGTGTGCTGCTGCTGCC 377
QY 304 TATGACCCACAGAGAGCTGGGAGAGCGACACAATGCTGGGAGGAGGCTGCGAGGCG 363
Db 378 TATGACCCACAGAGAGCTGGGAGAGCGACACAATGCTGGGAGGAGGCTGCGAGGCG 437
QY 364 GTGATCCCGCGGCTCTCATGGAACCTCTGAGCTCACAAGGGAGGAGGCTGCGAGGCG 423
Db 438 GTGATCCCGCGGCTCTCATGGAACCTCTGAGCTCACAAGGGAGGAGGCTGCGAGGCG 497
QY 424 CGGCGATGGGCGCTTCTGTCACATGCTTACCTAGAGATCTACAGAGAGAGTATTA 483
Db 498 CGGCGATGGGCGCTTCTGTCACATGCTTACCTAGAGATCTACAGAGAGAGTATTA 557
QY 484 GACCTCTGAGAGCTGCTTCCGAGAGCTGCTGTAATCCGAGAGAGCTGCGGGGAAATATC 543
Db 558 GACCTCTGAGAGCTGCTTCCGAGAGCTGCTGTAATCCGAGAGAGCTGCGGGGAAATATC 617
QY 544 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTTCAGCGGCACTTC 603
Db 618 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTTCAGCGGCACTTC 677
QY 604 CTGCCAGCCAGTGTGAAATCGAAGTGTAGGAGCCACCCGCTCAACAGGCTCTCTCCGCG 663
Db 678 CTGCCAGCCAGTGTGAAATCGAAGTGTAGGAGCCACCCGCTCAACAGGCTCTCTCCGCG 737
QY 664 AGTCATGCTGTGCTCTGCTCAAGGTGGAGCCAGCGGGAACGTTTGGGCGCCATTTCCGCG 723
Db 738 AGTCATGCTGTGCTCTGCTCAAGGTGGAGCCAGCGGGAACGTTTGGGCGCCATTTCCGCG 797
QY 724 CGAGAGGGAAATCTTACCTGATGACTTTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 783
Db 798 CGAGAGGGAAATCTTACCTGATGACTTTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 857
QY 784 AACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCTGTTGCTGCTGGG 843
Db 858 AACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCTGTTGCTGCTGGG 917
QY 844 AAAGTGTAGTGTGCTGCTCAAGTGTGAGTGTGCTGATTTTTCAGCGGCAAGCTC 903
Db 918 AAAGTGTAGTGTGCTGCTCAAGTGTGAGTGTGCTGATTTTTCAGCGGCAAGCTC 977
QY 904 ACTGCGCTATTGACAGACTCTCTGGGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAG 963
Db 978 ACTGCGCTATTGACAGACTCTCTGGGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAG 1037
QY 964 GCGCCCTGAGAGCGCTTCTACCTAGACACAGTCTCGGCACTCAACTTGTGCTGCGGCTCC 1023
Db 1038 GCGCCCTGAGAGCGCTTCTACCTAGACACAGTCTCGGCACTCAACTTGTGCTGCGGCTCC 1097
QY 1024 AAGGAGGTGATCAATTG 1040

```

```
Db 1098 AAGGAGGTGATCAATCG 1114
|||||
RESULT 9
US-09-595-684B-34
; Sequence 34, Application US/09595684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Ohaishi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Vaisberg, Eugene
; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; TITLE OF INVENTION: and purifying human kinesins
; FILE REFERENCE: cytopo36
; CURRENT APPLICATION NUMBER: US/09/595,684B
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 2097
; TYPE: DNA
; ORGANISM: Human
US-09-595-684B-34

Query Match 99.5%; Score 1035.4; DB 4; Length 2097;
Best Local Similarity 99.9%; Pred. No. 9.9e-315;
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GTCTCGTCTCGCTAGCAGATGGAGTACTCGTCTGTCACCTCCAGCTCCGCTAAGG 63
Db 100 GTCTCGTCTCGGTAAGCAAGATGGAGTACTCGTCTGTCACCTCCAGCTCCGCTAAGG 159
QY 64 GTGCGTGTGCGACTCGCGCCATTGTGGATGGAACACGCGGAGCAAGTGTATCCCTCTGT 123
Db 160 GTGCGTGTGCGACTCGCGCCATTGTGGATGGAACACGCGGAGCAAGTGTATCCCTCTGT 219
QY 124 GTGCGGGCATGACAGCTCTCTCTAGAGATTGCTAACTGGAGGACCAACCCAGGAGACT 183
Db 220 GTGCGGGCATGACAGCTCTCTCTAGAGATTGCTAACTGGAGGACCAACCCAGGAGACT 279
QY 184 CTCAATACAGTTTGATGCTCTCTAGAGATTGCTAACTGGAGGACCAACCCAGGAGACT 243
Db 280 CTCAATACAGTTTGATGCTCTCTAGAGATTGCTAACTGGAGGACCAACCCAGGAGACT 339
QY 244 GGTTCAGTGCAGCCCACTCTAAGGCATCTGCTGGAAGGGCAGAAATGCAAGTGTGCTTGGC 303
Db 340 GGTTCAGTGCAGCCCACTCTAAGGCATCTGCTGGAAGGGCAGAAATGCAAGTGTGCTTGGC 399
QY 304 TATGGACCCACAGAGCTGGGAAGACGACACAATGCTGGGCGAGCCCGAGCAACCTGGG 363
Db 400 TATGGACCCACAGAGCTGGGAAGACGACACAATGCTGGGCGAGCCCGAGCAACCTGGG 459
QY 364 GTGATCCGCGGGCTCTCATGAGCTCTCTGAGCTCAAGAGGAGAGGGTCCGAGGGC 423
Db 460 GTGATCCGCGGGCTCTCATGAGCTCTCTGAGCTCAAGAGGAGAGGGTCCGAGGGC 519
QY 424 CGGCCATGGGCCCTTTCTGTCACCATGCTTACTAGAGATCTACAGGAGAGAGGTATTA 483
Db 520 CGGCCATGGGCCCTTTCTGTCACCATGCTTACTAGAGATCTACAGGAGAGAGGTATTA 579
QY 484 GACCTCTGAGCCCTGCTTCCGAGAGCTCTGTTAATCCGAGAGACTGCCGGGGGAATATC 543
Db 580 GACCTCTGAGCCCTGCTTCCGAGAGCTCTGTTAATCCGAGAGACTGCCGGGGGAATATC 639
QY 544 CTGATTCGCGGGTCTCTCCAGAGGCCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTC 603
Db 640 CTGATTCGCGGGTCTCTCCAGAGGCCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTC 699

QY 604 CTGCCAGCCAGTGGAAATCGGACTGTAGGAGCCACCGGGCTCAACAGAGCGTCTCTCCGCG 663
Db 700 CTGCCAGCCAGTGGAAATCGGACTGTAGGAGCCACCGGGCTCAACAGAGCGTCTCTCCGCG 759
QY 664 AGTCATGCTGTCTCTGTTGTTCAAGGTGGACCAAGCGGGAACGTTTGGCCCCCATTTCCGCG 723
Db 760 AGTCATGCTGTCTCTGTTGTTCAAGGTGGACCAAGCGGGAACGTTTGGCCCCCATTTCCGCG 819
QY 724 CGAGAGGGAACCTCTACCTGATTGACTTGGCTGGGTTCAGAGGACCAACCGGCGCACAGGC 783
Db 820 CGAGAGGGAACCTCTACCTGATTGACTTGGCTGGGTTCAGAGGACCAACCGGCGCACAGGC 879
QY 784 AACAGGGGCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCTGGCG 843
Db 880 AACAGGGGCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCTGGCG 939
QY 844 AAAGTGTAGATCGCTGAATCAGGGCTCCCTCGTGTACTCTTATCGGACAGCAAGCTC 903
Db 940 AAAGTGTAGATCGCTGAATCAGGGCTCCCTCGTGTACTCTTATCGGACAGCAAGCTC 999
QY 904 ACTCGCTATTGCGAGACTCTCTGGGTGCTCAGCCACAGATATCTTATTCGCAAGCTC 963
Db 1000 ACTCGCTATTGCGAGACTCTCTGGGTGCTCAGCCACAGATATCTTATTCGCAAGCTC 1059
QY 964 GCCCTTGAGAGCGCTTCTACCTAGACACAGTCTCGGCACTCAACTTTGCTGCCAGGTCC 1023
Db 1060 GCCCTTGAGAGCGCTTCTACCTAGACACAGTCTCGGCACTCAACTTTGCTGCCAGGTCC 1119
QY 1024 AAGGAGGTGATCAATTG 1040
Db 1120 AAGGAGGTGATCAATCG 1136

RESULT 10
US-09-883-096-6
; Sequence 6, Application US/09883096
; Patent No. 6680369
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Craven, Andrew
; APPLICANT: Yu, Ming
; APPLICANT: Sakowicz, Roman
; APPLICANT: Patel, Umesh A.
; APPLICANT: Davies, Katherine A.
; TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE
; FILE REFERENCE: 020552-001410US
; CURRENT APPLICATION NUMBER: US/09/883,096
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 09/594,655
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1152
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HsKip3a
; OTHER INFORMATION: fragment
; OTHER INFORMATION: Nucleotide sequence of HsKip3a fragment used in
; OTHER INFORMATION: ATPase assay (Figure 5).
US-09-883-096-6

Query Match 13.1%; Score 136.6; DB 4; Length 1152;
Best Local Similarity 51.2%; Pred. No. 1.2e-32;
Matches 437; Conservative 0; Mismatches 389; Indels 27; Gaps 4;

QY 196 TTTGATGCTCTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCGAGTTTCAGTGCAG 255
Db 214 TTTGACCGGGTCTTTGGCGAGGCGGCCACCCAAAGAGCGTGTTCAGGACACACCGCCAC 273
QY 256 CCCATCCTAAGGCACCTTGCTGGAAGGGCAGAAATGCCAGTGTGCTTATGAGCCACCA 315
```

Db 274 AGGCTCTGGACAGCTTCTCTCAGGGCTACACTGCTCAGTGTGCTTACGGGCGCAC 333
Qy 316 GGAGCTGGAGAGCAGCACAATGCTGGGAGCCAGAGAACCTGGGGTGATGCCGGG 375
Db 334 GGGCTGGGAAGACACACCATGCTGGGAAGGGAGGAGCCCGGCATCATGTA--- 389
Qy 376 GCTCTATGGAACCTCTGAGCTCACAAGGAGGAGGTGCCGAGGGCGGCATGGGCC 435
Db 390 -----CCTGACACCGTGGAACTGTACAGGCGCTGGAGGCGCGCAGCAGAGAGCAC 444
Qy 436 CTTTCTGTCAACATGCTTACCTTAGAGATCTACAGGAAGATATTAGACCTCTCTGGAC 495
Db 445 TTGAGGTGCTCATCAGCTACAGGAGGTGTATATGACATCATCATCTCTGGAG 504
Qy 496 CTGCTTGGGAGACCTGTATATCCAGAAAGACTGCCGGGGAAATCTGTATTCGGGT 555
Db 505 C---CCAAGGGGCCCTTGCCATCCGAGAGGACCCCGACAAGGGGGTGGTGAAGA 561
Qy 556 CTCTCCAGAGCCCATCAGTCTTGTCTGATTTTGGAGGCACTTCTCTGCCAGCAT 615
Db 562 CTTTCTTTCCACAGCCAGCTCAGCCGAGAGCTGCTGGAGATATGACAGGGGAAAC 621
Qy 616 CGAAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCTCCGCACTCATGCTGT 675
Db 622 CGTAACCGCAGCAGCACCCCACTGATGCCAAGCGACTTCTCTCGCTCCCATGCCATC 681
Qy 676 CTCCTGGTCAAGGTGACAGCGGGAAAGCTTTGGCCCCATTTGGCCAG-----CGAGAG 729
Db 682 TTCCAGATCTTTGTGAAGCAGCAGGACCGGGTTCCAGGACTGACCCAGGCTGTCCAGGTG 741
Qy 730 GGAATACTTACTGATTAAGTCTGCTGGTTCAGAGGACACCGGGCAGAGCAACAG 789
Db 742 GCAAGATGAGCTGATGACTGGCTGAGCGGGGCAATCCAGCCCAATCCAGGAG 801
Qy 790 GGCCTTCGGCTAAAAGAGAGTGGAGCCATCAACACCTCCCTCTGTTGCTTGGGCAAGTG 849
Db 802 GGGAGCGCTGGGAGGGGCGCAACATCAACCGCTCTCTGCTGGCGCTCATCAACGTC 861
Qy 850 GTAGATCGCTGATCAGG-----CCCTCCCTCGGTGATCTTATTCGCAAC 900
Db 862 CTAATGCTTGGCGATCAAAAGGCGCAAGACCCATGTGCTTCCCTACCGGACAGCAA 921
Qy 901 CTCACCTCGCTATTCAGAGCTCTCTGGGTGCTCAGCCACAGTATCTTATTCGAAC 960
Db 922 CTGACCGCTGCTCAAGACTCCCTCGGGGCAACTGCGGCACAGTATGATCGTCC 981
Qy 961 ATTGCCCTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAG 1020
Db 982 ATCAGCCCTCCAGCTGACCTACGAGGACAGTACACACCTCAATATCGCGACCG 1041
Qy 1021 TCAAGGAGGTGA 1033
Db 1042 GCAAGGAGATCA 1054

RESULT 11

US-09-883-096-1
; Sequence 1, Application US/09883096
; Patent No. 6680369
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Craven, Andrew
; APPLICANT: Yu, Ming
; APPLICANT: Sakowicz, Roman
; APPLICANT: Patel, Umesh A.
; APPLICANT: Davies, Katherine A.
; TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE
; FILE REFERENCE: 020552-001410US
; CURRENT APPLICATION NUMBER: US/09/883,096
; PRIORITY FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 09/594,655
; FILING DATE: 2000-06-15

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4108
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Nucleic acid sequence of human kinesin motor
; OTHER INFORMATION: protein gene Hskip3a (Figure 1).
; OTHER INFORMATION: Description of Artificial Sequence: Hskip3a gene.
US-09-883-096-1

Query Match 13.1%; Score 136.6; DB 4; Length 4108;
Best Local Similarity 51.2%; Pred. No. 2.4e-32; Mismatches 389; Indels 27; Gaps 4;
Matches 437; Conservative 0;
Qy 136 TTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCAAGTTTCAGTGGC 255
Db 344 TTTGACCGGTCTTTGGCGAGCGGCCACCAACAGGAGGTGTTCCAGCACACCGCAC 403
Qy 256 CCATCTCTAAGCACTTGTGAGAGGAGAGTGCAGTGTGCTTGTCTATGACCCACA 315
Db 404 AGCGTCTCGACAGCTTCTCCAGGGCTACAACCTGCTCAGTGTGTTGCTTACGGGGCCAC 463
Qy 316 GGAGCTGGAGAGCAGCACAATGCTGGGAGCGCCAGAGCAACCTGGGGTGTATCCCGGG 375
Db 464 GGGGCTGGAGAGACACACCATGCTGGAGAGGGAGGGGACCCCGGATCATGTA---- 519
Qy 376 GCTCTATGAGACCTCTGCACTCAAGGGAGGAGGTGCGAGGGCGGCGCATGGGCC 435
Db 520 -----CCTGACCCAGCTGTACAGGCGCTGGAGGCCCGCCAGCAGGAGAAC 574
Qy 436 CTTTCTGTCAAGTGTCTTACCTAGAGATCTACAGGAGAGATATTAGACCTCTCTGAC 495
Db 575 TTCCAGGTGCTCATCAGCTACAGGAGGTGTATATGAACAGATCATGATCTCTGGAG 634
Qy 496 CTGCTTGGGAGACTGCTGTAATCGAGAGACTGCGGGGGAATATCTGATTCGGGT 555
Db 635 C---CCAAGGGGCCCTTGCCATCCGCGAGAACCCCGAAGGGGGTGGTGTGAAGA 691
Qy 556 CTCTCCAGAGCCCATCAGTGTGCTGATTTTGAGCGGCACTTCTGCGACCGCAT 615
Db 692 CTTTCTTCCACAGCCAGCTCAGCGAGCAGCTGTGGAGATACTCAGCAGGGGGAAC 751
Qy 616 CGAATCGAGCTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGCACTGATCTGTG 675
Db 752 CGTAACCGCAGCAGCACCCCACTGATGCCAAACGCACTTCTCTCCGCTCCCATGCCATC 811
Qy 676 CTCTGCTCAAGGTGAGCAGCGGGAACGTTTGGGCCCATTTGCGCAG-----CGAGAG 729
Db 812 TTCCAGATCTTTGAGCAGAGGACCGGGTTCAGGACTGACCCAGGCTGTCCAGGTG 871
Qy 730 GGAATACTCTA CTTGATGACTTGTGCTGGGTGAGAGCAACCGGCGCAGAGCAACAG 789
Db 872 GCAAGATGAGCTGATTTGACCTGGCTCAGAGCGGCAATCCAGCACCCATCGCAAG 931
Qy 790 GGCCTTCGCTAAAGAGAGTGGAGCCATCAACCTCCCTCTTGTCTTGGGCAAGTG 849
Db 932 GGGGAGCGGCTGCGGGAGGGGCCAACAATCAACCGCTCTCTGCTGGCGCTCATCAACGTC 991
Qy 850 GTAGATGCGCTGAATCAGG-----GCCTCCCTCGTGTACCTTATCGGAGACAGCAAG 900
Db 992 CTCATGCTTTGGCGGATGCAAGGGCGCAAGCCCATGTGCTCTACCGGAGACGAAA 1051
Qy 901 CTCACCTCGCTATTGACAGACTCTCTGGGTGGCTCAGCCCAAGPATCTTATTGCCAAC 960
Db 1052 CTGACCGCTGCTCAAGAGACTCTCTCGGGGCAACTGCGGCACAGTATGATCGCTGCC 1111
Qy 961 ATTGCCCTGAGAGAGCGTCTTACCTAGACAGACTCTCGGCACTCAACTTTGCTGCCAG 1020
Db 1112 ATCAGCCCTCCAGCCTGACCTACGAGGACAGTACAACACCTCAATATATCGCGACCG 1171
Qy 1021 TCAAGGAGGTGA 1033

```
Db 1172 GCCAAGGAGATCA 1184
|||||
RESULT 12
US-09-621-233-1
; Sequence 1, Application US/09621233
; Patent No. 6294371
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6294371el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1045
; CURRENT APPLICATION NUMBER: US/09/621.233
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 897
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (472)...(495)
; OTHER INFORMATION: n = a, c, t, or g
US-09-621-233-1

Query Match 12.3%; Score 127.8; DB 3; Length 897;
Best Local Similarity 50.1%; Pred. No. 5.9e-30;
Matches 429; Conservative 0; Mismatches 356; Indels 72; Gaps 3;

Qy 190 TACCAGTTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCAGGTTCA 249
|||
Db 106 TACCTGTTCAGTGGCTTTGACTTACCGCCACCCAGAGATGGTGTATCAGGCCACC 155
|||
Qy 250 GTGCAGGCCATCTTAAAGCACTTCTGGAAGGCGAGATGCCAGTGTGCTTGCCTATGGA 309
|||
Db 166 ACCAAGAGCTCATCGAGGGCGTCACTCAGGCTACAAATGCCACTGTCTTGCCTATGGC 225
|||
Qy 310 CCCACAGGAGCTGGGAAGACGCAACAATGCTGGGAGCCAGCAGCACTTGGGTTGATC 369
|||
Db 226 CCCACAGGCTGTGGGAAACCTACACATGCTGGGACACAGCCAGGAGCTTGCATCTAT 285
|||
Qy 370 CCGCGGCTCTCAGTGGACCTCTGAGCTCACAAGGAGGAGGAGTGCAGGAGGCGGCCA 429
|||
Db 286 GTTCAGACCTCAACGACCTTCTCCGTGCCATCGAGGAGACCAGCAATGA-----C 336
|||
Qy 430 TGGGCCCTTCTGTCAACCATGTCTTACCTAGAGATCTACCAGGAGAGGTATTAGACCTC 489
|||
Db 337 ATGAGTATGAGTCTCCATGCTCTACTTGGAGATCTACAAATGAGATGATCCGGACCTG 396
|||
Qy 490 CTGGAACCTGCTTCCGAGAGCTTGGTAATCCGAGAAGACTGCCGGGGGAATCTCTGATT 549
|||
Db 397 CTGAACCCCTCCCTGGGCTACCTGGAGCTCGGGAGGACTCTAAGGGGGTGTATCCAGGTG 456
|||
Qy 550 CCGGGTCTCTCCAGAGGCCATCAGTATGCTTGTGCTGATTTTGAGGGGCACTTCCCTGCCA 609
|||
Db 457 GCGGCAATACCGAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 516
|||
Qy 610 GCCAGTCAAAATCGGACTGTAGGAGCCACCGGCTCAACACAGAGCTCTCCCGCAGTCAAT 669
|||
Db 517 GGGAAACGGCAGAGGACCCAGGAGCCCAACCGGCGCCCAACAGAGCTCTCCGCTCCAC 576
|||
Qy 670 GCTGTGCTCTCTGTCAGGTGGACCGCGGGAACGTTTGGCCCAATTTGCCAGCGAGAG 729
|||
Db 577 GCGGGCGCGCTGTCA----- 592
|||
Qy 730 GGAATACTTACCTGATTGCTGGCTGGGTGAGGACAAACCGGCGCAGGCAACAAG 789
|||
Db 593 -----TGATCGACCTTGGCTGGCTCAGAGCGCGCCCTCGCAGACACAGAAATCGT 639
|||
Qy 790 GGCCCTTGGCTAAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGGGAAAGTG 849
|||
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2004, 19:47:28 ; Search time 4600.34 Seconds
(without alignments)
10701.074 Million cell updates/sec

Title: US-10-797-893-5
Perfect score: 1041
Sequence: 1 atgggtcgctgtcggtgaag.....ccaaggagtgatcaattga 1041

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1041	100.0	1041	6	AR210054 Sequence
2	1038.4	99.8	1464	6	AR210055 Sequence
3	1038	99.7	1115	6	AR210052 Sequence
4	1035.4	99.5	1538	6	AR210053 Sequence
5	1035.4	99.5	1998	9	BT007259 Homo sapi
6	1035.4	99.5	1998	12	BT007888 Synthetic
7	1035.4	99.5	2097	6	AR304057 Sequence
8	1035.4	99.5	2097	9	AB017430 Homo sapi
9	1035.4	99.5	2117	9	BC028155 Homo sapi
10	1035.4	99.5	2134	9	BC004352 Homo sapi
11	772.8	74.2	2086	10	BC003427 Mus muscu
12	678	65.1	151041	2	AC101919 Mus muscu
13	678	65.1	164759	2	AC101752 Mus muscu
14	650.8	62.5	297639	2	AC101908 Mus muscu
15	522.8	50.2	196674	10	AC102127 Mus muscu
16	462.2	44.4	2756	5	BC063896 Xenopus t
17	454.6	43.7	2565	5	BC043733 Xenopus l
18	454.6	43.7	2583	5	XL249841 Xenopus l
19	454.6	43.7	2640	5	BC073177 Xenopus l

20	438.6	42.1	2593	5	AF267850	Xenopus l
21	438.6	42.1	2661	5	XL249840	Xenopus l
22	438.6	42.1	2781	5	BC070549	Xenopus l
23	438.6	42.1	2819	5	AF267849	Xenopus l
24	355.4	34.1	174007	9	AC023831	Xenopus l
25	355.4	34.1	174363	9	AC009133	Homo sapi
26	352.2	33.8	3396	9	AB017333S2	Homo sapi
27	346.4	33.3	444	10	AF013119	Mus muscu
28	343.4	33.0	105485	9	HUAC002301	Homo sapi
29	278.4	26.7	68916	2	AC101927	Mus muscu
30	265	25.5	228783	2	AC094377	Rattus no
31	265	25.5	235806	2	AC133784	Rattus no
32	265	25.5	289504	2	AC123480	Rattus no
33	262.4	25.2	195987	10	AC122537	Mus muscu
34	262.4	25.2	226601	10	AC122863	Mus muscu
35	205.2	19.7	1982	3	AK114179	Ciona int
36	184.6	17.7	2095	6	AX833125	Sequence
37	184.6	17.7	2095	9	AK094619	Homo sapi
38	153.2	14.7	146415	5	AL929469	Zebrafish
39	153.2	14.7	176906	2	CR354587	Danio rer
40	151.8	14.6	3358	10	BC057614	Mus muscu
41	145.8	14.0	435	5	MSU34658	Morone saxa
42	145.2	13.9	44004	2	AC092583	Hylobates
43	143	13.7	183586	9	AC092582	Papio ham
44	136.6	13.1	1152	6	AR454169	Sequence
45	136.6	13.1	4108	6	AR454167	Sequence

ALIGNMENTS

RESULT 1
LOCUS AR210054 1041 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 5 from patent US 6387644.
ACCESSION AR210054
VERSION AR210054.1 GI:21512185.
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1041)
AUTHORS Beraud,C.
TITLE Motor proteins and methods for their use
JOURNAL Patent: US 6387644-A 5 14-MAY-2002;
FEATURES
source Location/Qualifiers
1..1041
/organism="unknown"
/mol_type="unassigned DNA"

Query Match	100.0%	Score 1041;	DB 6;	Length 1041;
Best Local Similarity	100.0%;	Pred. No. 1.9e-267;		
Matches 1041;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCAGCTCGCGTA	60	
Db	1	ATGGGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCAGCTCGCGTA	60	
QY	61	AGGGTGGCTGTGCGACTGCGGCATTTGTGGATGGAAACAGCGGGAGCAAGTGTATCCCCC	120	
Db	61	AGGGTGGCTGTGCGACTGCGGCATTTGTGGATGGAAACAGCGGGAGCAAGTGTATCCCCC	120	
QY	121	TGTGTGGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGACCAACCCAGGAG	180	
Db	121	TGTGTGGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGACCAACCCAGGAG	180	
QY	181	ACTCTCAATACACAGTTTCATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTAT	240	
Db	181	ACTCTCAATACACAGTTTCATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTAT	240	
QY	241	GCAGGTTAGTCAGCGCCCATCTTGTCTGGAAGGGCAGAAATCCAGTGTGCTT	300	

Db 241 GCAGGTTTCAGTCAGCCCATCTTAAGGCATTGCTGGAAGGCGAGAAATGCCAGTGTGCTT 300
 QY 301 GCCTATGAGACCCACAGAGCTGGGAAGAGCGCACAAATGCTGGGCGACCCAGAGCAACCT 360
 Db 301 GCCTATGAGACCCACAGAGCTGGGAAGAGCGCACAAATGCTGGGCGACCCAGAGCAACCT 360
 QY 361 GGGGTGATCCCGGGGCTCTCATGGACCTCCTGACCTCCTGACCTCACAGGAGAGGGTGGCAG 420
 Db 361 GGGGTGATCCCGGGGCTCTCATGGACCTCCTGACCTCCTGACCTCACAGGAGAGGGTGGCAG 420
 QY 421 GGCAGGCTATGGGCTCTCTTCTGTCACCATGCTCTTACCTAGAGATCTTACAGAGAGGTA 480
 Db 421 GGCAGGCTATGGGCTCTCTTCTGTCACCATGCTCTTACCTAGAGATCTTACAGAGAGGTA 480
 QY 481 TTAGACCTCTGAGACCTCTGCTGCTGGGAGACCTGGTAATCCGAGAAGACTGCCGGGGGAAT 540
 Db 481 TTAGACCTCTGAGACCTCTGCTGCTGGGAGACCTGGTAATCCGAGAAGACTGCCGGGGGAAT 540
 QY 541 ATCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGCGGCAC 600
 Db 541 ATCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGCGGCAC 600
 QY 601 TTCTGCGCAGCCAGTGGAAATCGGACTGTAGGAGCCACCGGCTCAACCGAGCGCTCCTCC 660
 Db 601 TTCTGCGCAGCCAGTGGAAATCGGACTGTAGGAGCCACCGGCTCAACCGAGCGCTCCTCC 660
 QY 661 CCGAGTCATGCTGCTCTGCTCAAGGTGGACCGAGCGGGAACGTTTGGCCCAATTTGCG 720
 Db 661 CCGAGTCATGCTGCTCTGCTCAAGGTGGACCGAGCGGGAACGTTTGGCCCAATTTGCG 720
 QY 721 CAGCGAGAGGGAATCTCTACCTGATTGCTGGTGGCTGAGAGCAACCGGCGCACA 780
 Db 721 CAGCGAGAGGGAATCTCTACCTGATTGCTGGTGGCTGAGAGCAACCGGCGCACA 780
 QY 781 GGCAACAGGGCTTGGCTAAAGAGAGTGGAGCCATCAACCTCCTGTTTGTCTG 840
 Db 781 GGCAACAGGGCTTGGCTAAAGAGAGTGGAGCCATCAACCTCCTGTTTGTCTG 840
 QY 841 GGCAAGTGTAGTGGCTGAATCAGGGCTCCTGCTGATACCTTATCGGAGAGCAAG 900
 Db 841 GGCAAGTGTAGTGGCTGAATCAGGGCTCCTGCTGATACCTTATCGGAGAGCAAG 900
 QY 901 CTCACCTCGCTATTTGAGAGACTCTCTGGGTGGCTCAGGCCCTCCCTGCTGATACCTTATTCGCAAC 960
 Db 901 CTCACCTCGCTATTTGAGAGACTCTCTGGGTGGCTCAGGCCCTCCCTGCTGATACCTTATTCGCAAC 960
 QY 961 ATTGCCCCCTGAGAGAGCTTCTACCTAGACAGTCTCGGACTCAACTTTGCTGCCAGG 1020
 Db 961 ATTGCCCCCTGAGAGAGCTTCTACCTAGACAGTCTCGGACTCAACTTTGCTGCCAGG 1020

RESULT 2
 AR210055
 LOCUS AR210055 1464 bp DNA linear PAT 20-JUN-2002
 DEFINITION Sequence 7 from patent US 6387644.
 ACCESSION AR210055
 VERSION AR210055.1 GI:21512186
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 UNCLASSIFIED.
 REFERENCE 1 (bases 1 to 1464)
 AUTHORS Beraud,C.
 TITLE Motor proteins and methods for their use
 JOURNAL Patent: US 6387644-A 7 14-MAY-2002;
 FEATURES Location/Qualifiers
 1..1464
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN

Query Match 99.8%; Score 1038.4; DB 6; Length 1464;
 Best Local Similarity 99.9%; Pred. No. 9.5e-267;
 Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGGGTCTCTGTCGGCTAAGCAAGATTGGAGCTACTCGTGTGCTACCTCCAGTCCGCTA 60
 Db 1 ATGGGTCTCTGTCGGCTAAGCAAGATTGGAGCTACTCGTGTGCTACCTCCAGTCCGCTA 60
 QY 61 AGGGTGGCTGTGCGGCTGCGGCTATTTCTGATGGAACAGCGGGGAGCAAGTATCCCCC 120
 Db 61 AGGGTGGCTGTGCGGCTGCGGCTATTTCTGATGGAACAGCGGGGAGCAAGTATCCCCC 120
 QY 121 TGTGTGCGGGGCTATGGACAGCTGCTCTTAGAGATTGCTAACTGAGGAAACCAACGAG 180
 Db 121 TGTGTGCGGGGCTATGGACAGCTGCTCTTAGAGATTGCTAACTGAGGAAACCAACGAG 180
 QY 181 ACTCTCAAAATACCAAGTTTGTATGCTTTCTATGGGAGAGAGTACTCTCAGCAGCATCTAT 240
 Db 181 ACTCTCAAAATACCAAGTTTGTATGCTTTCTATGGGAGAGAGTACTCTCAGCAGCATCTAT 240
 QY 241 GCAGTTTCAGTGCAGCCCATCTTAAGGCACCTTCTGGAAGGGGCAAGTGCAGTGTGCTT 300
 Db 241 GCAGTTTCAGTGCAGCCCATCTTAAGGCACCTTCTGGAAGGGGCAAGTGCAGTGTGCTT 300
 QY 301 GCCTATGGAACCCACAGGAGCTGGGAAGCGCACAAATGCTGGGCGAGCCCGAGCAACCT 360
 Db 301 GCCTATGGAACCCACAGGAGCTGGGAAGCGCACAAATGCTGGGCGAGCCCGAGCAACCT 360
 QY 361 GGGGTGATTCGGGCTCTCATGGAACCTCTGAGCTCAACAGGAGAGGGTGGCAG 420
 Db 361 GGGGTGATTCGGGCTCTCATGGAACCTCTGAGCTCAACAGGAGAGGGTGGCAG 420
 QY 421 GCGCGGCTATGGGCTTCTGTCACCATGCTTCTACCTAGAGATCTACAGAGAGAGTGA 480
 Db 421 GCGCGGCTATGGGCTTCTGTCACCATGCTTCTACCTAGAGATCTACAGAGAGAGTGA 480
 QY 481 TTAGACCTCTGGAACCTGCTTTCGGGAGACCTGTAATCCGAGAAGACTGCCGGGGGAAT 540
 Db 481 TTAGACCTCTGGAACCTGCTTTCGGGAGACCTGTAATCCGAGAAGACTGCCGGGGGAAT 540
 QY 541 ATCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTTGGAGCGGCAC 600
 Db 541 ATCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTTGGAGCGGCAC 600
 QY 601 TTCTGCGCAGCCAGTGGAAATCGGACTGTAGGAGCCACCGGCTCAACCGAGCGCTCCTCC 660
 Db 601 TTCTGCGCAGCCAGTGGAAATCGGACTGTAGGAGCCACCGGCTCAACCGAGCGCTCCTCC 660
 QY 661 CCGAGTCATGCTGCTCTGCTCAAGGTGGACCGAGCGGGAACGTTTGGCCCAATTTGCG 720
 Db 661 CCGAGTCATGCTGCTCTGCTCAAGGTGGACCGAGCGGGAACGTTTGGCCCAATTTGCG 720
 QY 721 CAGCGAGAGGGAATCTCTACCTGATTGCTGGTGGCTCAGAGGCAACCGGCGCACA 780
 Db 721 CAGCGAGAGGGAATCTCTACCTGATTGCTGGTGGCTCAGAGGCAACCGGCGCACA 780
 QY 781 GGCAACAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCTGTTTGTCTG 840
 Db 781 GGCAACAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCTGTTTGTCTG 840
 QY 841 GGCAAGTGTAGTGGCTGAATCAGGGCTCCTGCTGATACCTTATCGGAGAGCAAG 900
 Db 841 GGCAAGTGTAGTGGCTGAATCAGGGCTCCTGCTGATACCTTATCGGAGAGCAAG 900
 QY 901 CTCACCTCGCTATTTGAGAGACTCTCTGGGTGGCTCAGGCCCTCCCTTATTCGCAAC 960
 Db 901 CTCACCTCGCTATTTGAGAGACTCTCTGGGTGGCTCAGGCCCTCCCTTATTCGCAAC 960
 QY 961 ATTGCCCCCTGAGAGAGCTTCTACCTAGACAGTCTCGGACTCAACTTTGCTGCCAGG 1020
 Db 961 ATTGCCCCCTGAGAGAGCTTCTACCTAGACAGTCTCGGACTCAACTTTGCTGCCAGG 1020

```
QY 1021 TCCAGGAGGTGATCAATTG 1040
Db 1021 TCCAGGAGGTGATCAATCG 1040

RESULT 3
AR210052
LOCUS AR210052 1115 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 1 from patent US 6387644.
ACCESSION AR210052
VERSION AR210052.1 GI:21512182
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1115)
AUTHORS Beraud,C.
TITLE Motor proteins and methods for their use
JOURNAL Patent: US 6387644-A 1 14-MAY-2002;
FEATURES
    Location/Qualifiers
        source
            1..1115
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN
    Query Match 99.7%; Score 1038; DB 6; Length 1115;
    Best Local Similarity 100.0%; Pred. No. 1.2e-266;
    Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGCGTAAGG 63
Db 78 GGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGCGTAAGG 137

QY 64 GTGGCTGTGCGCTGCGGCAATTGTGGATGGAACAGCGGGAGCAAGTATCCCCCTGT 123
Db 138 GTGGCTGTGCGCTGCGGCAATTGTGGATGGAACAGCGGGAGCAAGTATCCCCCTGT 197

QY 124 GTGCGGGGATGACAGAGCTGCTCTCTAGAGATTGCTTAAGTGGAGAACACACAGGAGACT 183
Db 198 GTGCGGGGATGACAGAGCTGCTCTCTAGAGATTGCTTAAGTGGAGAACACACAGGAGACT 257

QY 184 CTCAATACCAAGTTGATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTATGCA 243
Db 258 CTCAATACCAAGTTGATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTATGCA 317

QY 244 GGTTCAGTGCAGCCATCTTAGGCATCTGCTGGAAGGCAAGATGCCAGTGTGCTTGC 303
Db 318 GGTTCAGTGCAGCCATCTTAGGCATCTGCTGGAAGGCAAGATGCCAGTGTGCTTGC 377

QY 304 TATGACCCACAGGAGCTGGGAAGAGCGACACAAATGCTGGGAGCCCGAGCAACTGGG 363
Db 378 TATGACCCACAGGAGCTGGGAAGAGCGACACAAATGCTGGGAGCCCGAGCAACTGGG 437

QY 364 GTGATCCCGCGGCTCTCATGACCTCTGACAGTCTCAAGGAGAGAGGCTCCGAGGCG 423
Db 438 GTGATCCCGCGGCTCTCATGACCTCTGACAGTCTCAAGGAGAGAGGCTCCGAGGCG 497

QY 424 CGGCCATGCGGCCCTTCTGTACCAATGCTTACCTAGATCTACAGGAGAGGTATTA 483
Db 498 CGGCCATGCGGCCCTTCTGTACCAATGCTTACCTAGATCTACAGGAGAGGTATTA 557

QY 484 GACCTCTCGGACCTGCTTCGGGAGACCTGGTAATCCGAGAGAGCTCGCGGGGAATATC 543
Db 558 GACCTCTCGGACCTGCTTCGGGAGACCTGGTAATCCGAGAGAGCTCGCGGGGAATATC 617

QY 544 CTGATTCCGGGTCTCTCCCAAGGCCCATCAGTAGCTTTGCTGATTTTGGAGCGCACTTC 603
Db 618 CTGATTCCGGGTCTCTCCCAAGGCCCATCAGTAGCTTTGCTGATTTTGGAGCGCACTTC 677

QY 604 CTGCAGCCAGTGCAGAAATCGGACTGTAGGAGCCACCGGCTCAACAGGCTCTCCCGC 663
Db 678 CTGCAGCCAGTGCAGAAATCGGACTGTAGGAGCCACCGGCTCAACAGGCTCTCCCGC 737
```

```
QY 664 AGTCATGCTGTCTCTCTGCTCAAGGTGGACACGCGGGAAGCTTTGGCCCATTTTCGCCAG 723
Db 738 AGTCATGCTGTCTCTCTGCTCAAGGTGGACACGCGGGAAGCTTTGGCCCATTTTCGCCAG 797

QY 724 CGAGAGGGAAAACTCTACTGATTGACTTTGGCTGGGTCTAGAGGACAAACCGCGCACAGGC 783
Db 798 CGAGAGGGAAAACTCTACTGATTGACTTTGGCTGGGTCTAGAGGACAAACCGCGCACAGGC 857

QY 784 AACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTTGTCTTGGGC 843
Db 858 AACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTTGTCTTGGGC 917

QY 844 AAAGTGGTAGATGCGCTGAATCAGGGCCTCCCTCGTGTACCTTATCGGACACCAAGCTC 903
Db 918 AAAGTGGTAGATGCGCTGAATCAGGGCCTCCCTCGTGTACCTTATCGGACACCAAGCTC 977

QY 904 ACTCGCTTATTCAGAGCTCTCTGCGTGGCTCAGGCCACAGTATCTTATGCAACATT 963
Db 978 ACTCGCTTATTCAGAGCTCTCTGCGTGGCTCAGGCCACAGTATCTTATGCAACATT 1037

QY 964 GCCCCTGAGAGAGCTTCTACCTAGACACAGTCTCGGCACTCAACTTTGCTGCCAGTCC 1023
Db 1038 GCCCCTGAGAGAGCTTCTACCTAGACACAGTCTCGGCACTCAACTTTGCTGCCAGTCC 1097

QY 1024 AAGGAGGTGATCAATTGA 1041
Db 1098 AAGGAGGTGATCAATTGA 1115

RESULT 4
AR210053
LOCUS AR210053 1538 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 3 from patent US 6387644.
ACCESSION AR210053
VERSION AR210053.1 GI:21512183
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1538)
AUTHORS Beraud,C.
TITLE Motor proteins and methods for their use
JOURNAL Patent: US 6387644-A 3 14-MAY-2002;
FEATURES
    Location/Qualifiers
        source
            1..1538
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN
    Query Match 99.5%; Score 1035.4; DB 6; Length 1538;
    Best Local Similarity 99.9%; Pred. No. 6.1e-266;
    Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGTCTCTGTGCGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGCGTAAGG 63
Db 78 GGTCTCTGTGCGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGCGTAAGG 137

QY 64 GTGGCTGTGCGACTCGCGCCATTTGTGATGGAACAGCGGGAGCAAGTATCCCCCTGT 123
Db 138 GTGGCTGTGCGACTCGCGCCATTTGTGATGGAACAGCGGGAGCAAGTATCCCCCTGT 197

QY 124 GTGCGGGGATGACAGAGCTGCTCTCTAGAGATTGCTTAAGTGGAGAACACACAGGAGACT 183
Db 198 GTGCGGGGATGACAGAGCTGCTCTCTAGAGATTGCTTAAGTGGAGAACACACAGGAGACT 257

QY 184 CTCAATACCAAGTTGATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTATGCA 243
Db 258 CTCAATACCAAGTTGATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTATGCA 317

QY 244 GGTTCAGTGCAGCCATCTTAGGCATCTGCTGGAAGGCAAGATGCCAGTGTGCTTGC 303
Db 318 GGTTCAGTGCAGCCATCTTAGGCATCTGCTGGAAGGCAAGATGCCAGTGTGCTTGC 377
```


QY 364 GTGATCCCGGGCTCTCATGGACCTCTGTGAGCTCAAGGGAGAGGGTGGCGAGGGC 423
 Db |||||
 QY 436 GTGATCCCGGGCTCTCATGGACCTCTGTGAGCTCAAGGGAGAGGGTGGCGAGGGC 495
 Db |||||
 QY 424 CGGCCATGGCCCTTCTGTGACCATCTTACCTAGAGTCTACACAGGAGAGGTATTA 483
 Db |||||
 QY 496 CGGCCATGGCCCTTCTGTGACCATCTTACCTAGAGTCTACACAGGAGAGGTATTA 555
 Db |||||
 QY 484 GACCTCTCGGACCTCTGTGACCATCTTACCTAGAGTCTACACAGGAGAGGTATTA 543
 Db |||||
 QY 556 GACCTCTCGGACCTCTGTGACCATCTTACCTAGAGTCTACACAGGAGAGGTATTA 615
 Db |||||
 QY 544 CTGATTCGGGGTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGTAGCGGCACTTC 603
 Db |||||
 QY 616 CTGATTCGGGGTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGTAGCGGCACTTC 675
 Db |||||
 QY 604 CTGCCACGAGTCCAGATCGGAGTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCCGC 663
 Db |||||
 QY 676 CTGCCACGAGTCCAGATCGGAGTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCCGC 735
 Db |||||
 QY 664 AGTCATCTGTGCTCTCTGGTCAAGGTGGACCGAGCGGAAGTGTGGCCCATTTTCGCCAG 723
 Db |||||
 QY 736 AGTCATCTGTGCTCTCTGGTCAAGGTGGACCGAGCGGAAGTGTGGCCCATTTTCGCCAG 795
 Db |||||
 QY 724 CGAGAGGAGAACTCTACCTGATTTGCTGCTGGTCAAGGACCAACCGCGCACAGGC 783
 Db |||||
 QY 796 CGAGAGGAGAACTCTACCTGATTTGCTGCTGGTCAAGGACCAACCGCGCACAGGC 855
 Db |||||
 QY 784 AACAGGGGCTCTGGCTAAAGAGTGGAGCCATCAACCTCCCTGTTTGTCTGGC 843
 Db |||||
 QY 856 AACAGGGGCTCTGGCTAAAGAGTGGAGCCATCAACCTCCCTGTTTGTCTGGC 915
 Db |||||
 QY 844 AAAGTGTAGATGCGCTGAATCAGGGCTCCCTGCTGATCTTATCGGACAGCAAGCTC 903
 Db |||||
 QY 916 AAAGTGTAGATGCGCTGAATCAGGGCTCCCTGCTGATCTTATCGGACAGCAAGCTC 975
 Db |||||
 QY 904 ACTGCGCTATTGAGGACTCTCTGGTGGCTCAGCCCAAGTATCTTATGCAACATT 963
 Db |||||
 QY 976 ACTGCGCTATTGAGGACTCTCTGGTGGCTCAGCCCAAGTATCTTATGCAACATT 1035
 Db |||||
 QY 964 GCCCTGAGAGCGCTCTACCTAGACACAGTCTCCGACCTCACTTTGCTGCGAGTCC 1023
 Db |||||
 QY 1036 GCCCTGAGAGCGCTCTACCTAGACACAGTCTCCGACCTCACTTTGCTGCGAGTCC 1095
 Db |||||
 QY 1024 AAGGAGGTGATCAATTG 1040
 Db |||||
 QY 1096 AAGGAGGTGATCAATCG 1112
 Db |||||

RESULT 6
 BT007888
 LOCUS 1998 bp mRNA linear SYN 13-MAY-2003
 DEFINITION Synthetic construct Homo sapiens kinesin-like 4 mRNA, partial cds.
 ACCESSION BT007888
 VERSION BT007888.1 GI:30584614
 KEYWORDS FLI CDNA.
 SOURCE synthetic construct
 ORGANISM artificial sequences.
 REFERENCE 1 (bases 1 to 1998)
 AUTHORS Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y., Phelan,M., and Farmer,A.
 TITLE Cloning of human full-length CDSs in BD Creator(TM) System Donor vector
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1998)
 AUTHORS Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y., Phelan,M., and Farmer,A.
 TITLE Direct Submission
 JOURNAL Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303, USA

COMMENT

This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two forms: with and without stop-codon (to allow fusion with C-terminal tag). The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the SalI and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'ACC' after SalI site and before 'ATG' to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame. Clone distribution: <http://bioinfo.clontech.com/orfclones>.

FEATURES

source

1..1998
 /organism="synthetic construct"
 /mol_type="mRNA"
 /db_xref="taxon:32630"
 /clone="GH0024511.0"
 /clone_lib="BD Creator(TM) CDS Library derived from MOC collection"
 /lab_host="DH5alpha T1 resistant"
 /note="Vector: pDNR-Dual"

CDS

1..>1998
 /note="Mutations: 1997:Stop->Leu"
 /codon_start=1
 /transl_table=11
 /product="Homo sapiens kinesin-like 4"
 /protein_id="AAP36560.1"
 /db_xref="GI:30584615"
 /translation="MAAGSTQQRREMAAASAAISGAGRCLSKIGATRRPPARV
 RVARURPFVDGTAGADPPVGRGMSCLSIANWRNHQTLTKYQFADYGERSTQD
 IYAGVQPIILHLLGONASVLAAYGTAGKTHTMLGSPQPGVPPALMDLLQTLRE
 AGAEPFWALSVTMYLEIYQEKVLDLDPASGDVIRDCRGNILIPGLSQKPISSF
 ADPERHFLPASNRVTGRLNORSSRAVLVVKVDRLAPRFRQEKLYLIDIA
 GSDNRRTNGKLRKESGAINTSFLVGVKVDALNQLGPRVPRDSKLRLLQDSLG
 GSAHSILIANIAPERPYLDTUSALNPAARSKVINRPTNESLQPHALGPKLSQKE
 LIGPPAKARGPPEEIEGPEPMAAPASAKSLQKLSMDPAMLEKRLSLDRLL
 ASQSGGAPLLSTPKRVMVMTVEEKDLEIRLTKQKEAKMLAQAKAEKNHC
 PTMLRSLHRTVTGPKLKKAVVMPQLIQEQAASPAEIHILNKGRKKRKLSDLA
 EPEKAEDCWEIQISPELLAHGRQKILDLINEGSARDLRSQIRGPKKQLIVGREL
 HGPFQSVDELVERVEGITGQKMFSLKILGLAAGQRCASL"

ORIGIN

Query Match 99.5%; Score 1035.4; DB 12; Length 1998;
 Best Local Similarity 99.9%; Pred. No. 6.1e-266;
 Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCTGCTCCACCTCCAGCTCGGCTAAGG 63
 Db |||||
 QY 76 GGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCTGCTCCACCTCCAGCTCGGCTAAGG 135
 Db |||||
 QY 64 GTGCTGTGCGCTGTGCGCCATTGTGGATGGAAACAGCGGAGCAAGTGATCCCCCTGT 123
 Db |||||
 QY 136 GTGCTGTGCGCTGTGCGCCATTGTGGATGGAAACAGCGGAGCAAGTGATCCCCCTGT 195
 Db |||||
 QY 124 GTGCGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACT 183
 Db |||||
 QY 196 GTGCGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACT 255
 Db |||||
 QY 184 CTCAAATACAGTTTTCATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 243
 Db |||||
 QY 256 CTCAAATACAGTTTTCATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 315
 Db |||||
 QY 244 GGTTCAGTGCAGGCCATCTTAAGGCATCTTGTGGAAGGGCAGAAATGCCAGTGTGCTGGC 303
 Db |||||
 QY 316 GGTTCAGTGCAGGCCATCTTAAGGCATCTTGTGGAAGGGCAGAAATGCCAGTGTGCTGGC 375
 Db |||||
 QY 304 TATGGACCCACAGGAGCTGGGAAGACGACACAATCTGCGGACGCCCAGAGCAACCTGGG 363
 Db |||||
 QY 376 TATGGACCCACAGGAGCTGGGAAGACGACACAATCTGCGGACGCCCAGAGCAACCTGGG 435
 Db |||||
 QY 364 GTGATCCCGGGGCTCTCATGGACCTCTGTCAGCTCACAAGGAGAGGGTGGCGAGGGC 423
 Db |||||
 QY 436 GTGATCCCGGGGCTCTCATGGACCTCTGTCAGCTCACAAGGAGAGGGTGGCGAGGGC 495
 Db |||||

QY	424	CGGCCATGGGCCCTTCTGTCAACATGCTTACCTAGAGATCTACCGAGAGAGTATTA	483
DB	496	CGGCCATGGGCCCTTCTGTCAACATGCTTACCTAGAGATCTACCGAGAGAGTATTA	555
QY	484	GACCTCTCGGACCTTCTCGGAGACCTTGGTAATCCGAGAGAGCTCGCGGGGGAATATC	543
DB	556	GACCTCTCGGACCTTCTCGGAGACCTTGGTAATCCGAGAGAGCTCGCGGGGGAATATC	615
QY	544	CTGATTCGGGGTCTCTCCAGAAAGCCCATCACTAGTCTTTGTGATTTTGAAGCGGCACTTC	603
DB	616	CTGATTCGGGGTCTCTCCAGAAAGCCCATCACTAGTCTTTGTGATTTTGAAGCGGCACTTC	675
QY	604	CTGCCAGCCAGTCGAATTCGGAATCTGAGAGCCACCCGGCTCAACAGCGGCTCTCCCGC	663
DB	676	CTGCCAGCCAGTCGAATTCGGAATCTGAGAGCCACCCGGCTCAACAGCGGCTCTCCCGC	735
QY	664	AGTCATGCTGCTCTCTGCTCAAGGTGAGCGGAGCAACGTTTGGCCCATTTTCGCCAG	723
DB	736	AGTCATGCTGCTCTCTGCTCAAGGTGAGCGGAGCAACGTTTGGCCCATTTTCGCCAG	795
QY	724	CGAGAGGGAATACTTACCTGATTAATGACTTGGCTGGGTGAGAGCAACCGCGGCAAGGC	783
DB	796	CGAGAGGGAATACTTACCTGATTAATGACTTGGCTGGGTGAGAGCAACCGCGGCAAGGC	855
QY	784	ACAAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCCTGTTTCTCTGGGC	843
DB	856	ACAAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCCTGTTTCTCTGGGC	915
QY	844	AAAGTGTAGTACGCTGATCAGGGCTTCCCTGCTGTGATCTTATCGGGAGCAAGCTC	903
DB	916	AAAGTGTAGTACGCTGATCAGGGCTTCCCTGCTGTGATCTTATCGGGAGCAAGCTC	975
QY	904	ACTGCGCTATTCAGAGACTCTCTGGGTGGCTCAGGCCACAGTATCTTATTCGCAACATT	963
DB	976	ACTGCGCTATTCAGAGACTCTCTGGGTGGCTCAGGCCACAGTATCTTATTCGCAACATT	1035
QY	964	GCCCTGAGAGAGCTTCTACCTAGACAGCTCCGCACTCAACTTTGCTGCCAGGTCC	1023
DB	1036	GCCCTGAGAGAGCTTCTACCTAGACAGCTCCGCACTCAACTTTGCTGCCAGGTCC	1095
QY	1024	AAGGAGGTGATCAATTG 1040	
DB	1096	AAGGAGGTGATCAATCG 1112	
<p>RESULT 7</p> <p>AR304057</p> <p>LOCUS</p> <p>DEFINITION</p> <p>Sequence 34 from patent US 6544766.</p> <p>ACCESSION</p> <p>AR304057</p> <p>VERSION</p> <p>AR304057.1</p> <p>GI:31692955</p> <p>KEYWORDS</p> <p>Unknown.</p> <p>SOURCE</p> <p>Unknown.</p> <p>ORGANISM</p> <p>Unclassified.</p> <p>REFERENCE</p> <p>1 (bases 1 to 2097)</p> <p>Beraud,C., Ohashi,C., Sakowicz,R., Vaisberg,S., Wood,K. and Yu,M.</p> <p>AUTHORS</p> <p>Human kinases and methods of producing and purifying human</p> <p>TITLE</p> <p>kinases</p> <p>JOURNAL</p> <p>Patent: US 6544766-A 34 08-APR-2003;</p> <p>FEATURES</p> <p>Location/Qualifiers</p> <p>source</p> <p>1..2097</p> <p>/organism="unknown"</p> <p>/mol_type="genomic DNA"</p> <p>ORIGIN</p> <p>Query Match 99.5%; Score 1035.4; DB 6; Length 2097;</p> <p>Best Local Similarity 99.9%; Pred. No. 6.1e-266;</p> <p>Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;</p>			
QY	4	GGTCGCTGCGGCTAAGCAAGATTGGAGCTACTCGTCGCTCCACCTCCAGCTCGCTAAGG	63
DB	100	GGTCGCTGCGGCTAAGCAAGATTGGAGCTACTCGTCGCTCCACCTCCAGCTCGCTAAGG	159

QY	64	GTGGCTGTGCAGCTGCGGCCAATTTGTGATGGAACAGCGGAGCAAGTGATCCGCCCTGT	123
DB	160	GTGGCTGTGCAGCTGCGGCCAATTTGTGATGGAACAGCGGAGCAAGTGATCCGCCCTGT	219
QY	124	GTGCGGGGCATGGAAGAGCTGCTCTAGAGATTGCTAACTGGAGGAAACCAACAGAGACT	183
DB	220	GTGCGGGGCATGGAAGAGCTGCTCTAGAGATTGCTAACTGGAGGAAACCAACAGAGACT	279
QY	184	CTCAAAATACCAAGTTTGTATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA	243
DB	280	CTCAAAATACCAAGTTTGTATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA	339
QY	244	GGTTCACTGAGCCATCTTAAGCACTTGTCTGGAAGGCGCAAGATGCCAGTGTGCTGCC	303
DB	340	GGTTCACTGAGCCATCTTAAGCACTTGTCTGGAAGGCGCAAGATGCCAGTGTGCTGCC	399
QY	304	TATGGACCCACAGAGCTGGGAAGACGACACAATGCTGGGAGCCACAGAGCAACCTGGG	363
DB	400	TATGGACCCACAGAGCTGGGAAGACGACACAATGCTGGGAGCCACAGAGCAACCTGGG	459
QY	364	GTGATCCCGCGGCTCTCATGGAACCTCTGAGCTCAACAGGAGAGGCTGCCAGAGGC	423
DB	460	GTGATCCCGCGGCTCTCATGGAACCTCTGAGCTCAACAGGAGAGGCTGCCAGAGGC	519
QY	424	CGGCCATGGGCTTCTGTCACCAATGCTTACCTAGAGATCTACCAAGAGAGGTAATTA	483
DB	520	CGGCCATGGGCTTCTGTCACCAATGCTTACCTAGAGATCTACCAAGAGAGGTAATTA	579
QY	484	GACCTCTCGGACCTTCTCGGAGACCTGGTAATCCAGAGAGCTGCCGGGGAATATC	543
DB	580	GACCTCTCGGACCTTCTCGGAGACCTGGTAATCCAGAGAGCTGCCGGGGAATATC	639
QY	544	CTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGTGATTTTGAAGGCGACATTC	603
DB	640	CTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGTGATTTTGAAGGCGACATTC	699
QY	604	CTGCCAGCCAGTCGAATTCGAGCTGTAGAGCCACCCGGCTCAACAGCGCTCTCCCGC	663
DB	700	CTGCCAGCCAGTCGAATTCGAGCTGTAGAGCCACCCGGCTCAACAGCGCTCTCCCGC	759
QY	664	AGTCATGCTGCTCTCTGTCAGAGTGGAACAGCGGGAACGTTTGGCCCCATTTGCCAG	723
DB	760	AGTCATGCTGCTCTCTGTCAGAGTGGAACAGCGGGAACGTTTGGCCCCATTTGCCAG	819
QY	724	CGAGAGGGAATCTCTACTGATTGCTTGGCTGAGAGCAACCGGCGCACAGGC	783
DB	820	CGAGAGGGAATCTCTACTGATTGCTTGGCTGAGAGCAACCGGCGCACAGGC	879
QY	784	AACAAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACCTCTGTTTGTCTGGGC	843
DB	880	AACAAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACCTCTGTTTGTCTGGGC	939
QY	844	AAAGTGTAGATGCGCTGAATCAGGGCTCTGCTGCTGATCTTATCGGAGACAGAGCTC	903
DB	940	AAAGTGTAGATGCGCTGAATCAGGGCTCTGCTGCTGATCTTATCGGAGACAGAGCTC	999
QY	904	ACTCGCTTATTCAGAGCTCTCTGGGTGGCTCAGCCACAGTATCTTATTCGCAACATT	963
DB	1000	ACTCGCTTATTCAGAGCTCTCTGGGTGGCTCAGCCACAGTATCTTATTCGCAACATT	1059
QY	964	GCCCTCTAGAGAGCGCTTCTACCTAGACAGCTCTCCGCACTCAACTTTGCTGCCAGTCC	1023
DB	1060	GCCCTCTAGAGAGCGCTTCTACCTAGACAGCTCTCCGCACTCAACTTTGCTGCCAGTCC	1119
QY	1024	AAGGAGGTGATCAATTG 1040	
DB	1120	AAGGAGGTGATCAATCG 1136	

RESULT 8

AB017430

LOCUS

AB017430 2097 bp mRNA linear PRI 06-MAR-1999

REFERENCE
AUTHORS

1 (bases 1 to 2117)
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Scheinman C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schenker C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L.,
Sheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S.,
Carinci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,
Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J.,
McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettunen M., Madan A., Rodriguez S.,
Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y.,
Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S., Krzywinski M.I., Skalska U., Smallus D.E.,
Schnerch A., Schein J.E., Jones S.J. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

TITLE

human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
PUBMED

12477932

REFERENCE
AUTHORS

2 (bases 1 to 2117)
Strausberg R.
Direct Submission
Submitted (08-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nigr.nih.gov
Akhter N., Ayele K., Beckstrom-Sternberg S.M., Benjamin B.,
Blakesley R.W., Bouffard G.G., Breen K., Brinkley C., Brooks S.,
Dietrich N.L., Grant R., Guan X., Gupta J., Haghighi P.,
Hansen N., Ho S.-L., Karlins E., Kwong P., Laric P., Legaspi R.,
Maduro Q.L., Masiello C., Maskeri B., Mastriani S.D., McCloskey J.C.,
McDowell J., Pearson R., Stancipop S., Thomas P.J., Touchman J.W.,
Tsurgueon C., Vogt J.L., Walker M.A., Wetherby K.D., Wiggins L.,
Young A., Zhang L.-H. and Green E.D.

REMARK
COMMENT

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Series: IRAC Plate: 62 Row: 1 Column: 3
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.
Location/Qualifiers
1. .2117
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:40049 IMAGE:5241557"
/tissue_type="Brain, fetal, whole pooled"
/clone_lib="NIH MGC_121"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .2117
/gene="KIF22"
/note="synonym: KID, OBP, OBP-1, OBP-2, KNSL4"
/db_xref="LocusID:3835"
/db_xref="MIM:603213"
6. .2003
/gene="KIF22"

FEATURES
source

1. .2117
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:40049 IMAGE:5241557"
/tissue_type="Brain, fetal, whole pooled"
/clone_lib="NIH MGC_121"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .2117
/gene="KIF22"
/note="synonym: KID, OBP, OBP-1, OBP-2, KNSL4"
/db_xref="LocusID:3835"
/db_xref="MIM:603213"
6. .2003
/gene="KIF22"

gene

1. .2117
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:40049 IMAGE:5241557"
/tissue_type="Brain, fetal, whole pooled"
/clone_lib="NIH MGC_121"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .2117
/gene="KIF22"
/note="synonym: KID, OBP, OBP-1, OBP-2, KNSL4"
/db_xref="LocusID:3835"
/db_xref="MIM:603213"
6. .2003
/gene="KIF22"

CDS

1. .2117
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:40049 IMAGE:5241557"
/tissue_type="Brain, fetal, whole pooled"
/clone_lib="NIH MGC_121"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .2117
/gene="KIF22"
/note="synonym: KID, OBP, OBP-1, OBP-2, KNSL4"
/db_xref="LocusID:3835"
/db_xref="MIM:603213"
6. .2003
/gene="KIF22"

/codon_start=1
/product="kinesin family member 22"
/protein_id="AAH28155.1"
/db_xref="GI:20380447"
/db_xref="LocusID:3835"
/db_xref="MIM:603213"
/translation="MAAGGTCQORREMAAASAAISAGRCRLSKIGATRRPPARV
RVVRLRPFVDTGASDPCVRGMDSCSLEIANRNHOETIKQFDYFGERSTQDD
IYAGVDPILRLHLEGQNASVLYPTGAGTKHTLMSPEQPGVPIRALMDLLOITRE
EGASRFPALSVTMSYLEIYQEKVLDLDDPAGDLVIREDCRGNILIPGLSQPTSSF
ADFERHFLPASRNTTQATRNQSRSSHAVLLVKVDORERLAPPROEGKLYLIDLA
GSENNRTNGKRLKSGSANTSLFVLGVKVDALNOGLPRVPRVDSKLRLLQDSLG
GSAHLLIANTAPERREYLDTVSALNFAARSKVINRPTNESLQPHALGPVKLSQKE
LIGPEAKRAGPEEIEIGSPMAAPASAKLSPLQKSSMDPAMFLERLRLDLL
ASQSGAPLLSTPKRRLVMTVEKDLERIKTKQKLEAKMLAKAEBENHC
PTMLRPLSHRTVTGAKPLKAVVMPLOIQPAAAPNAIHLKNNKGRKGLSDAL
EPEKAECDWELQISPELLAHGROKILDLNLESGARDLSRLQIPKPKAQLIVGWREL
HGFPSQVEDLERVSGITGKQWESFLKANILGLAAGRCGAS"

ORIGIN

Query Match 99.5%; Score 1035.4; DB 9; Length 2117;
Best Local Similarity 99.9%; Pred. No. 6.1e-266;
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 GGTGCTGTCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCTAAGG 63
DB 81 GGTGCTGTCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCTAAGG 140
QY 64 GTGGCTGTGCACTGCGGCCATTTGTGATGGAACAGCGGGAGCAAGTGTATCCCTCTGT 123
DB 141 GTGGCTGTGCACTGCGGCCATTTGTGATGGAACAGCGGGAGCAAGTGTATCCCTCTGT 200
QY 124 GTGGGGGCGATGACAGCTGCTCTCTAGAGATTGCTAACTGAGAGAACACACAGAGACT 183
DB 201 GTGGGGGCGATGACAGCTGCTCTCTAGAGATTGCTAACTGAGAGAACACACAGAGACT 260
QY 184 CTCAAAATACCAAGTTGATGCTCTATGGGAGAGAGTACTCTAGAGAGACATCTATGCA 243
DB 261 CTCAAAATACCAAGTTGATGCTCTATGGGAGAGAGTACTCTAGAGAGACATCTATGCA 320
QY 244 GGTTCAGTGTGACGCCATCTTAAGGCACTTGTGGAAGGGCAAGATGCGAGTGTGCTGCC 303
DB 321 GGTTCAGTGTGACGCCATCTTAAGGCACTTGTGGAAGGGCAAGATGCGAGTGTGCTGCC 380
QY 304 TATGACCCACAGAGCTGGGAGACACACACATGCTGGGAGCCACACAGACACTGGG 363
DB 381 TATGACCCACAGAGCTGGGAGACACACATGCTGGGAGCCACACAGACACTGGG 440
QY 364 GTGATCCCGGGCTCTCATGACCTCTGAGCTCAACAGGAGAGGGTGGCGAGGGC 423
DB 441 GTGATCCCGGGCTCTCATGACCTCTGAGCTCAACAGGAGAGGGTGGCGAGGGC 500
QY 424 CGGCGATGGGCCCTTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGAAAGTATTA 483
DB 501 CGGCGATGGGCCCTTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGAAAGTATTA 560
QY 484 GACCTCTGTGACCCCTGCTTGGGAGACTGTGATCCGAGAGACTGCGGGGGAATATC 543
DB 561 GACCTCTGTGACCCCTGCTTGGGAGACTGTGATCCGAGAGACTGCGGGGGAATATC 620
QY 544 CTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGTGATTTTGTAGCGGCATCTC 603
DB 621 CTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGTGATTTTGTAGCGGCATCTC 680
QY 604 CTGCGACCGAGTCCAAATCGGACTGTAGAGGCCACCCGGCTCAACCGAGCGTCTCTCCGC 663
DB 681 CTGCGACCGAGTCCAAATCGGACTGTAGAGGCCACCCGGCTCAACCGAGCGTCTCTCCGC 740
QY 664 AGTCATGCTGTGCTCTGCTCAAGCTGACCGGGGAAAGCTTTGGCCCAATTTGGCCAG 723
DB 741 AGTCATGCTGTGCTCTGCTCAAGCTGACCGGGGAAAGCTTTGGCCCAATTTGGCCAG 800
QY 724 CGAGAGGGAAACTCTTACCTGATTGACTTGGCTGGGTGAGAGGAAACCGCGCGCACAGGC 783

```
Db      801  CGAGAGGGAACCTTACCTGATTTGCTTGGCTGAGGACACCGCGCAGGC 860
QY      784  AACAGGGCTTCGGCTAAAGAGAGTGGAGGCATCAACACCTCCCTGTTGTCCTGGGC 843
Db      861  AACAGGGCTTCGGCTAAAGAGAGTGGAGGCATCAACACCTCCCTGTTGTCCTGGGC 920
QY      844  AAGTGGTAGATCGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGACAGCAAGCTC 903
Db      921  AAGTGGTAGATCGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGACAGCAAGCTC 980
QY      904  ACTCGCTATTGAGAGACTCTTGGGTGGCTAGCCACAGATATCTTATGCAACATT 963
Db      981  ACTCGCTATTGAGAGACTCTTGGGTGGCTAGCCACAGATATCTTATGCAACATT 1040
QY      964  GCCCTTGAGAGCGCTTCTACCTAGACACAGTCTCGCACTCAACTTGTGCGAGGTCC 1023
Db      1041  GCCCTTGAGAGCGCTTCTACCTAGACACAGTCTCGCACTCAACTTGTGCGAGGTCC 1100
QY      1024  AAGGAGGTGATCAATTG 1040
Db      1101  AAGGAGGTGATCAATCG 1117
```

```
RESULT 10
LOCUS   BC004352
DEFINITION Homo sapiens kinesin family member 22, mRNA (CDNA clone MGC:1573
IMAGE:3535435), complete cds.
```

```
ACCESSION BC004352
VERSION   BC004352.1 GI:13279307
KEYWORDS MGC.
```

```
SOURCE   Homo sapiens
ORGANISM Homo sapiens (human)
```

```
REFERENCE 1 (bases 1 to 2134)
AUTHORS  Krausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
```

```
  Klausner,R.D., Collins,F.S., Wagner,L., Shemen,C.M., Schuler,G.D.,
  Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
  Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
  Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
  Stapleton,M., Soares,M.B., Bonaldo,M.P., Caevant,T.L.,
  Scheetz,T.B., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
  Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
  Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
  McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
  Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
  Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
  Fahy,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
  Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
  Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
  Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
  Butterfield,Y.S., Krzyzanski,M.I., Skalska,U., Smalun,D.E.,
  Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  Generation and initial analysis of more than 15,000 full-length
  human and mouse cDNA sequences
```

```
TITLE    Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL  12477932
PUBMED   2 (bases 1 to 2134)
AUTHORS  Strausberg,R.
```

```
TITLE    Direct Submission
JOURNAL  Submitted (01-MAR-2001) National Institutes of Health, Mammalian
  Gene Collection (MGC), Cancer Genomics Office, National Cancer
  Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
  USA
```

```
REMARK   NIH-MGC Project URL: http://mgi.nci.nih.gov
COMMENT  Contact: MGC help desk
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: DCTD/DTF
  CDNA Library Preparation: Rubin Laboratory
  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Genome Sequence Centre,
```

```
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim Macdonald, Anara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
Duane Smalun, Jeff Stott, Miranda Tsai, George Yang, Jacque
Schein, Asim Siddiqui, Rob Holt, Marco Marra.
```

```
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 8 Row: f Column: 1
```

```
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6453817.
```

FEATURES

```
Location/Qualifiers
1..2134
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="MGC:1573 IMAGE:3535435"
  /tissue_type="Lung, small cell carcinoma"
  /clone_lib="NIH MGC_7"
  /lab_host="DH10B-R"
  /note="Vector: pOTB7"
1..2134
  /gene="KIF22"
  /note="synonyms: KID, OBP, OBP-1, OBP-2, KNSL4"
  /db_xref="LocusID:3835"
  /db_xref="MIM:603213"
20..2017
  /gene="KIF22"
  /codon_start=1
  /product="kinesin family member 22"
  /protein_id="AAH04352.1"
  /db_xref="GI:13279308"
  /db_xref="LocusID:3835"
  /db_xref="MIM:603213"
```

CDS

```
RVAVRLRFVDTAGASPPCVRGMDSCSLKLEIANRWNETKLYQFDAYGERSTQDD
IYAGSVQPIRLHLLGQNASVLAYGTGAGKTHTWLGSPEQGVIPRALMDLLQLTRE
EGAEGRPWALSVTMSYLYEQKVDLLDPSGLVIREDCRGNLIPLGSKPISFF
ADFERHFLPASNRVTGATRLNQRSRSHAVLLVKVDORERLAPRQREGKLYLIDLA
GSEDNRRRTNGKLRLKESGAINTSFLVLGKVDALNOGLPRVYRDSKLTLLQDSLG
CSAHSILIANTAPERFYLDTVSALNFAARKEVINRPTNESLQPHALGPKVLSQKE
LIGPEAKRANGPEEEIGSEPFMAAPASASOKLSPLKLSMDPAMLRLLSLDKLL
ASQSGQAPLJSTPKRERMLMTVEEKDLEIRLKTQKELEAKMLAQAKBEKENHC
PTMLRPLSHRTVTGAKPKCAVMPLOEQASPNABEHLKNGKRRKRLKESLDAL
EPEKAEDCWELQISPELLAHGRQKILDLNLSGARDLRLSLQRIQPKKQLIVGWREL
HGPFQSVEDLERVEIGTGKQMSFLKANILGLAAGRCGAS"
```

ORIGIN

```
Query Match      99.5%; Score 1035.4; DB 9; Length 2134;
Best Local Similarity 99.9%; Pred. No: 6.1e-266;
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      4  GTCGCTGTCGGCTAAGCAAGATTGGAGCTACTGCTGCTCCACCTCCAGCTCGCTAAGG 63
Db      95  GGTGCTGTCGGCTAAGCAAGATTGGAGCTACTGCTGCTCCACCTCCAGCTCGCTAAGG 154
QY      64  GTGGCTGTGCGACTCGGCGCATTTGTGATGGAACAGCGGGAGCAAGTGTATCCCCCTGT 123
Db      155  GTGGCTGTGCGACTCGGCGCATTTGTGATGGAACAGCGGGAGCAAGTGTATCCCCCTGT 214
QY      124  GTGCGGGCATGACGACGCTGCTCTTAGAGATTGCTTAAGAGAAACCAACAGGAGACT 183
Db      215  GTGCGGGCATGACGACGCTGCTCTTAGAGATTGCTTAAGAGAAACCAACAGGAGACT 274
QY      184  CTCGAATACCAAGTTTGTATGCGCTTCTTAGGGGAGAGAGTACTCAGCAGACATCTATGCA 243
Db      275  CTCGAATACCAAGTTTGTATGCGCTTCTTAGGGGAGAGAGTACTCAGCAGACATCTATGCA 334
```

QY 244 GGTTCAGTGCAGCCCATCTAAGGCATCTGCTGGAAGGCGAGAAATGCCAGATGCTGCTTGC 303
Db 335 GTTTCAGTGCAGCCCATCTAAGGCATCTGCTGGAAGGCGAGAAATGCCAGATGCTGCTTGC 394
QY 304 TATGAGCCCAAGGAGCTGGGAAGACGACACAAATGCTGGGAGCCCAAGAGCAACTGGG 363
Db 395 TATGAGCCCAAGGAGCTGGGAAGACGACACAAATGCTGGGAGCCCAAGAGCAACTGGG 454
QY 364 GTGATCCCGCGGCTCTATGACCTCTGACCTCTGACCTCAAGGGAGGAGTCCGAGGCG 423
Db 495 GTGATCCCGCGGCTCTATGACCTCTGACCTCTGACCTCAAGGGAGGAGTCCGAGGCG 514
QY 424 CGGCATCGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACACGAGAGAGTATTA 483
Db 515 CGGCATCGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACACGAGAGAGTATTA 574
QY 484 GACCTCTGGACCTCTGCTGGGAGACTGTAATCCGAGAGAGCTGCCGGGGGAATATC 543
Db 575 GACCTCTGGACCTCTGCTGGGAGACTGTAATCCGAGAGAGCTGCCGGGGGAATATC 634
QY 544 CTGATCCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGAGGGGCACTTC 603
Db 635 CTGATCCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGAGGGGCACTTC 694
QY 604 CTGCAGGCAAGTGAATCGGACTCTAGGAGCCACCGGCTCAACGAGCGTCTCTCCGCG 663
Db 695 CTGCAGGCAAGTGAATCGGACTCTAGGAGCCACCGGCTCAACGAGCGTCTCTCCGCG 754
QY 664 AGTCATGCTGCTCTCTGCTCAAGGTGACCGAGCGGAACTGTTGGCCCCCAATTCGCCAG 723
Db 755 AGTCATGCTGCTCTCTGCTCAAGGTGACCGAGCGGAACTGTTGGCCCCCAATTCGCCAG 814
QY 724 CGAGGGGAAACTTACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 783
Db 815 CGAGGGGAAACTTACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 874
QY 784 AACAGGCGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACTCTGTTGCTGCTGCTGCT 843
Db 875 AACAGGCGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACTCTGTTGCTGCTGCTGCT 934
QY 844 AAGTGTGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 903
Db 935 AAGTGTGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 994
QY 904 ACTGCGCTATTGCGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 963
Db 995 ACTGCGCTATTGCGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1054
QY 964 GCGCTCTGAGAGCGCTTCTACTAGACACAGTCTCCGCACTCAACTTTGCTGCGAGGTC 1023
Db 1055 GCGCTCTGAGAGCGCTTCTACTAGACACAGTCTCCGCACTCAACTTTGCTGCGAGGTC 1114
QY 1024 AAGGAGGTGATCAATTG 1040
Db 1115 AAGGAGGTGATCAATCG 1131

RESULT 11
BC003427
LOCUS BC003427
DEFINITION Mus musculus kinesin family member 22, mRNA (cdna clone MGC:6456
IMAGE:2615715), complete cds.
ACCESSION BC003427
VERSION BC003427.1 GI:13097359
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2086)
REFERENCE Strausberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schneitz, F.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smal, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2086)
Strausberg, R.
Direct Submission
Submitted (20-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amgbc@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAC Plate: 5 Row: m Column: 7
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 21704181.

FEATURES

Location/Qualifiers
1..2086
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clones="MGC:6456 IMAGE:2615715"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal carcinoma, 5 month old virgin mouse."
/clone_lib="NCI CGAP_Mam6"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
1..2086
/gene="Kif22"
/note="synonym: Kid"
/db_xref="LocusID:110033"
/db_xref="MGI:109233"
20..2002
/gene="Kif22"
/codon_start=1
/product="kinesin family member 22"
/protein_id="AAH03427.1"
/db_xref="GI:13097360"
/db_xref="LocusID:110033"
/db_xref="MGI:109233"

gene

CDS

VQILRLHLEQNASVLAYGTGAGTKTHMLGSPQGVIPRALMDLLOLAREESAG
RPMVSVAMSYLRIYEDKVLDDLPASGDIIVIREDKRGNILIPCLTKPKITSFSDPEQ
HFLPASNRHAGATRLNORSRSHAVLLVVDORERLTPRQREGKYLIDLADSEEN
RRTNOCIRLKEGAINLTVLGVKVDALNOGLPRIPYRDSKLRLLOSLGSAHS
ILJANTAPERFYQDITISALNFTARSKEVINRPTNESLOPHALA PVKLSQKLELGS
EAKAKGPESSTGSPSTNAAPASOKSLQKLNMDPAMLENLLSMLERLLGSGS
QGPLNTPKRRMLVLMKVVEENKLEIRLUMKOKLEAKVLAQAPDPREKENTPTI
LOPPASVGTVAFLKAVVMPQRIQOKRESSNQIQLLKKPKRKLPEPSEAEVAK
DEYWEVQISPELLAHGRKLLDLNLEGSARELSLQRIQOKKALIVGWRELHGFPS
EVEDLEQVEGISGKQVESFLKANLLSLAASQHSGPS"

ORIGIN

Query Match 74.2%; Score 772.8; DB 10; Length 2086;
Best Local Similarity 84.3%; Pred. No. 1.2e-195;
Matches 870; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

QY	6	TCGCTGTCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCTGAAGGT	65
DB	82	TGCGCGCTGTGTAAAGCAAGAGGGCTTGGCCGGCGTCCGCTCTAGCTCGAGTAAGGT	141
QY	66	GGCTGTGCGACTCGCGCCATTGTGGATGAAACAGCGGAGCAAGTGATCCCCCTGTGT	125
DB	142	GGCTGTGCGCTGGCCCGTTTATGATGGAGAGACAGAGCGAGCTCCCTGTGT	201
QY	126	GGCGGGCAGTGGAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACACACAGAGACTCT	185
DB	202	CCGAGCCATAGACAGCTGCTCTCTTGAATGGCTAACTGGAAGAAATACACAGAGAGCT	261
QY	186	CAATACCACTTTGATGCTCTCTATGGGAGAGAGTACTCAGCAGACATCTATGAGG	245
DB	262	CAATATACGTTTGTATGCTCTTATGGCAGAGAGCACTCAGCAGGAAGTCTATGTAGG	321
QY	246	TTCACTGTCAGCCATCTTAAGGCACTTGCTGGAAGGGCAGAACTGCCTGTCTTGCTTA	305
DB	322	TTCACTGTCAGCCATCTTAAGGCACTTGCTGGAAGGGCAGAACTGCCTGTCTTGCTTA	381
QY	306	TGACCCACAGAGAGCTGGGAAGACGACACAATGCTGGGAGCCGACAGCAACTGGGGT	365
DB	382	TGGCGCTACTTGGGCGAGGAAGACACACAATGCTGGGAGCCGACAGCAACTGGAGT	441
QY	366	GATCCCGGGCTCTCATGACCTCTGAGCTCAGAGGAGGAGGAGGAGGAGGAGGAGG	425
DB	442	GATTCCTCGGGCTCTCATGACCTCTGAGCTCAGAGGAGGAGGAGGAGGAGGAGG	501
QY	426	GCCATGGCGCTTCTGTCTACCATGCTTACCTAGAGATCTACACAGAGAGGATATTAGA	485
DB	502	GCCATGGAGCTTCTGTCTGTCTATGCTTATGAGATCTACAGAAAGGATATTAGA	561
QY	486	CCTCTGGACCTCTGCTTCGGGAGACCTGGAATCCGAGAGACTGCGGGGGAAATATCCT	545
DB	562	CCTCTGGATCTCTGATCAGGAGACCTCGTGATCCGCGAGACTGCGGAGGAACATCCT	621
QY	546	GATTCGGGTCTCTCCAGAGGCCATCAGTACTGTTGCTGATTTGAGCGGCACTTCCT	605
DB	622	GATCCAGGCGCTCACAAGAGGCCATCACCACTCTCTGACTTCGAGCAGCACTTCCT	681
QY	606	GCCAGCAGTCGAAATCGGACTGAGAGCCACCGGCTCAACAGGCTCTCCCGCAG	665
DB	682	TCCAGCAGTAGAATTCGAGCTGAGAGGCCACCGGCTTAAACAGGCTCTTCCCGTAG	741
QY	666	TCATGCTGTCTCTCTGCTCAAGTGGACAGCGGGAACGTTTGGCCCGCAATTCGCCAGG	725
DB	742	TCACGCACTGCTCTTGTCAAGTGTAGATCAGCGTGAACGTTGACTCCATTTCCGCGAG	801
QY	726	AGAGGAAACTCTACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	785
DB	802	GGAAGGAAACTCTACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	861
QY	786	CAAGGCGCTTCGCTTAAAGAGAGTGGAGCCATCAACCTCCTCTGCTGCTGCTGCTG	845
DB	862	CAAGGCGCTTCGCTTAAAGAGAGTGGAGCCATCAACCTCCTCTGCTGCTGCTGCTG	921
QY	846	AGTGTGTAGTGCCTGAATCAGGCGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTG	905

DB	922	AGTGTGTGATGATTAACAGGCGCTCCCTCGCATACCATACCGGAGACAGCAAGCTCAC	981
QY	906	TCGCTATTTCAGGAGCTCTCTGGTGGCTCAGCCACAGTATCTTATTCACCAATGCG	965
DB	982	TCGCTGTGCTCAGGAGCTCTCTGGGAGCTCAGCTCATAGCATCTTATTCACCAATGCG	1041
QY	966	CCCTGAGAGAGCTTCTTACCTAGACAGCTCTCGGCACTCAACTTTCTGTCAGGTCACAA	1025
DB	1042	TCTTGAAGAGCTTTTACCAAGATCAATTTAGCATTCGATTCCTCTGCTAGTCCAA	1101
QY	1026	GGAGGTGATCAA 1037	
DB	1102	GGAGGTGATCAA 1113	

RESULT 12	
AC101919/c	
LOCUS	
DEFINITION	AC101919 151041 bp DNA linear HTG 14-MAR-2003
ACCESSION	Mus musculus clone RP24-158A6, WORKING DRAFT SEQUENCE, 4 unordered
VERSION	AC101919.4 GI:28951338
KEYWORDS	HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE	Birren, B., Nusbaum, C. and Lander, E.
JOURNAL	Mus musculus, clone RP24-158A6
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 151041)
REFERENCE	Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
AUTHORS	Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,
REFERENCE	Brown, A., Camarata, J., Campoliano, A., Chang, J., Chazaro, B.,
AUTHORS	Choepl, Y., Collangelo, M., Collins, S., Collamore, A., Cook, A.,
REFERENCE	Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
AUTHORS	Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
REFERENCE	Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
AUTHORS	Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
REFERENCE	Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K.,
AUTHORS	Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
REFERENCE	MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
AUTHORS	McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrin, J.,
REFERENCE	Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,
AUTHORS	Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
REFERENCE	Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
AUTHORS	Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
REFERENCE	Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
AUTHORS	Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
REFERENCE	Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
AUTHORS	Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliou, H.,
REFERENCE	Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
AUTHORS	Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE	Direct Submission
JOURNAL	Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
REFERENCE	Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS	3 (bases 1 to 151041)
REFERENCE	Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
AUTHORS	Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
REFERENCE	Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepl, Y.,
AUTHORS	Collamore, A., Cooke, A., Cooke, P., Corum, B., DeArellano, K., Faro, S.,
REFERENCE	Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
AUTHORS	Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
REFERENCE	Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
AUTHORS	Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
REFERENCE	Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
AUTHORS	Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
REFERENCE	Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
AUTHORS	Meldrin, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J.,
REFERENCE	Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
AUTHORS	O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,

Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Rhoman, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (14-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 14, 2003 this sequence version replaced gi:28631351.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu
 Project Information
 Center project name: L17745
 Center clone name: 158_A_6

Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 150558 bases at least Q40
 Consensus quality: 150654 bases at least Q30
 Consensus quality: 150687 bases at least Q20
 Insert size: 151000; agarose-fp
 Insert size: 150741; sum-of-contigs
 Quality coverage: 11.8 in Q20 bases; agarose-fp
 Quality coverage: 11.8 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 33517: contig of 33517 bp in length
 33617: gap of 100 bp
 33618 42736: contig of 9119 bp in length
 42737 42836: gap of 100 bp
 42837 84394: contig of 41558 bp in length
 84395 84494: gap of 100 bp
 84495 151041: contig of 66547 bp in length.

Location/Qualifiers
 1..151041
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="RP24-158A6"
 /clone_lib="RP24-158A6 Male Mouse BAC"
 misc_feature
 1..33517
 /note="assembly_fragment
 clone_end:SP6
 vector_side:left"
 misc_feature
 33618..42736
 /note="assembly_fragment"
 42837..84394
 /note="assembly_fragment"
 84495..151041
 /note="assembly_fragment
 clone_end:T7
 vector_side:right"

ORIGIN
 Query Match 65.1%; Score 678; DB 2; Length 151041;
 Best Local Similarity 80.0%; Pred. No. 2.8e-170;
 Matches 824; Conservative 0; Mismatches 200; Indels 6; Gaps 2;

12 TCGGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCAGCTCGCGTAAGGGTGGCTGT 71
 65699 TTGTATAAGCAAGGTAGGATTTGGCTGCTGCTTCTAGCTGGAGTAAGGGTGGCTAT 65640
 72 GCGACTGCGGCAATTTGTGGATGGAAAGCGGGAGCAAGTATCCCTCTGTGTGCGGG 131
 65639 ATAAATTGTCCTGTTTGTGGATGAAGACAGAGCAAGAGGACACCCCTGTGTACGAGC 65580
 132 CATGACAGCTGCTCTCTAGAGATTTGCTAACTGGAGGACACACAGGAGACTCTCAATA 191
 65579 CATGACAGCTGCTCTCTTGAAGTGGCTTAACCTGGAGGAAATACCAGGAGACTATCAAATA 65520
 192 CCAGTTTCATGCTCTTATGCGGAGAGAGTACTCAGCAGGACATCTATGACAGGTTTCACT 251
 65519 TCAGTTTATGCTTCCATGGAGAGAGAGACACTCAGTGGGACATCTATGTAGGTTTACT 65460
 252 GCAGCCCATCTAAGGCACTTGTGGAGGGGAGAAATCCAGTGTGCTTGTGCTATGGACC 311
 65459 ACAGCCCATCTAAGGAACTTGTGGAGGGGAGAAATCCAGTGTGCTTGTGCTATGGACC 65400
 312 CACAGGAGCTGGAGAGACGACACAACTGCTGGGAGGCCACAGCAACCTGGGGTATCCC 371
 65399 GACTGGGAGGGGAGAA--CACACAATGTTGGGAGGCCACAGCAACCTGGAGTGAATCCC 65342
 372 GCGGCTCTCATGAGACCTCTCTGAGCTCACAGGAGGAGGAGGTCGCGAGGGCGCCATG 431
 65341 TCAGACTCTCCAGACCTCTCTGACCTAACAGGGAAGAAATGCTGAGGGCTGGCCATG 65282
 432 GGCCCTTTCTGTACCACTGTCTTACCTAGAGATCTACAGGAGAAAGGTATTAGACTCTCT 491
 65281 GGACATTTCTGTGTTGTTGTTAGTCTCTTAACAGGAAATAGTATTAGACTCTCT 65222
 492 GGACCTCTCTGGGAGACCTGTTAATCGAGAGACTCGCGGGAATATCTCTGATCCC 551
 65221 GATCTCTGATCAGAGACCTTAGTGTCCGAGAGGCTGACAGGAGAAATCCTGATCCC 65162
 552 GGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGCGGCACTTCTCCCGAGC 611
 65161 AGGCTCTACACAGAGCCCATCAGTAGTTTCTCTGACTTCGAGAGAGCACTTCTCTTAGT 65102
 612 CAGTCGAATCGGACTGTAGAGGCCACCCGGCTCAACCAGGCTCTCCGAGCTATGC 671
 65101 CAGTAGAATCGAGTTGTAGGAGCTCTCTGTTAAACGAGGCTCTTCCGCTAGTCTATGC 65042
 672 TGTGCTCTCTGCTCAAGGTGGACACGCGGGAAGTTTGGCCCATTTCCGACGAGAGGG 731
 65041 AGCGCTCTTGGTCAAGGTAGATCAGCATGAACGTTTGACTCCATTTCCGCGAGTGGAGG 64982
 732 AAAACTCTACCTGATTGACTTGGCTGGGTGAGAGGACAAACC---GGCGCACAGGCAACA 787
 64981 AAAACTCTACCTTATTGATTTGGCTGGTTTCCAGAGGACAAACCGTCAGTCGACAGGCAACC 64922
 788 AGGCGCTTCCGCTAAAGAGAGTGAGCCATCAACACCTCCCTGTTTGTCTTGGGCAAG 847
 64921 AGGCGATTCTGCTCAAGAGAGTGAGGCCATCAACACCTCCCTCTTGTGTACTTGGGCAAG 64862
 848 TGGTAGATCGCTGAATCAGGCGCTCCCTCTCGTGTACTTATCGGACAGCAAGCTCACTC 907
 64861 TAGTGATGCTAATAACCAAGGCTCCCTCTCATATACATACAGGACAGCAGCTCACTC 64802
 908 GCCTATTGACGAGCTCTCTGGGTGGCTCAGCCACAGTATCTTATTCGCAATTTGCC 967
 64801 GTCTGTGCAAGACTCTCTGAGAGGCTCAGCTCATAGCATCTCTCATTTGTGAGCAATTTGCTC 64742
 968 CTGAGAGAGCTTCTTACCTAGACAGTCTCCGCACTCAACTTTCTGCTGCCAGGTCGAGG 1027
 64741 CTGAAAGATGTTTTTACCTGGATACATCTCAGCAATTAACCTTCACTGTAGGTCGAGG 64682
 1028 AGGTGATCAA 1037
 64681 AGGTGATCAA 64672

RESULT 13

AC101752/c
 LOCUS 164759 bp DNA linear ROD 23-APR-2004
 DEFINITION Mus musculus chromosome 1, clone RP24-344C18, complete sequence.
 ACCESSION AC101752
 VERSION
 KEYWORDS
 SOURCE HTG.
 ORGANISM Mus musculus (house mouse)

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS

1 (bases 1 to 164759)
 Birren, B., Nusbaum, C. and Lander, E.

TITLE

Mus musculus chromosome 1, clone RP24-344C18

REFERENCE

2 (bases 1 to 164759)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,

AUTHORS

Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,
 Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
 Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,
 Cooke, P., DeArelano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
 Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Kamat, A., Karatas, A., Kells, C., Lakocque, K.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
 MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,
 Meneus, L., Mihova, T., Menga, V., Murphy, T., Naylor, J., Nguyen, C.,
 Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Triglio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Direct Submission
 Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

3 (bases 1 to 164759)

AUTHORS

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
 Collamore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K.,
 Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
 Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., MacLean, C.,
 Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meldrim, J., Meneus, L., Mihova, T., Menga, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
 Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
 Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Direct Submission
 Submitted (17-MAR-2004) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

4 (bases 1 to 164759)

AUTHORS

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, M., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V.,
 Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J.,
 Choepel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B.,
 DeArelano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L.,
 Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D.,
 Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,

Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
 Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
 Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R.,
 MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,
 McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Menga, V.,
 Murphy, T., Naylor, J., Nguyen, C., O'Neill, D., Oliver, J., Peterson, K.,
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
 Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R.,
 Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
 Zimmer, A. and Zody, M.

TITLE

Direct Submission

JOURNAL

Submitted (23-APR-2004) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Apr 23, 2004 this sequence version replaced gi:45504330.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

REFERENCE

----- Genome Center

AUTHORS

Center: Whitehead Institute/MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@broad.mit.edu
 ----- Project Information
 Center project name: L17344
 Center clone name: 344_C_18

FEATURES

source
 Location/Qualifiers
 1..164759
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="1"
 /map="1"
 /clone="RP24-344C18"
 /clone_lib="RPC1-24 Male Mouse BAC"
 repeat_region 1..1655
 misc_feature 1..4
 /rpt_family="L1_MM"
 /note="clone boundary
 clone end:SP6
 site:Mbol"
 repeat_region 1767..1787
 /rpt_family="AT_rich"
 repeat_region 2115..2157
 /rpt_family="AT_rich"
 repeat_region 2236..2258
 /rpt_family="AT_rich"
 repeat_region 3252..3420
 /rpt_family="Lx7"
 repeat_region 3464..3588
 /rpt_family="Lx7"
 repeat_region 3605..3666
 /rpt_family="CT-rich"
 repeat_region 3673..3731
 /rpt_family="(TTC)n"
 repeat_region 3732..3911
 /rpt_family="(TTC)n"
 repeat_region 4000..4180
 /rpt_family="CT-rich"
 repeat_region 4202..4363
 /rpt_family="Lx7"
 repeat_region 4529..4765
 /rpt_family="MTE"
 repeat_region 7234..7442
 /rpt_family="B3"
 repeat_region 8445..8509
 /rpt_family="T-rich"
 repeat_region complement(8512..8591)
 /rpt_family="MLT11"


```

repeat_region 9336..9632
/rpt_family="MLT13"
complement(9800..11396)
/rpt_family="L1_MN"
complement(11397..17527)
/rpt_family="L1F"
complement(18235..18777)
/rpt_family="Lx"
repeat_region 18775..18890
/rpt_family="Lx"
repeat_region 19260..19949
/rpt_family="RMER19A"
repeat_region 20118..20297
/rpt_family="MLT1D"
complement(20386..20639)
/rpt_family="Lx2"
repeat_region 20634..20839
/rpt_family="Lx4"
repeat_region 20876..21793
/rpt_family="RLTR13D"
repeat_region 21799..21907
/rpt_family="RMER13A"
repeat_region 21918..22433
/rpt_family="Lx4"
repeat_region 22591..22651
/rpt_family="Lx4"
repeat_region 22733..24486
/rpt_family="L1_MN"
repeat_region 24494..24850
/rpt_family="L1_MN"
repeat_region 27139..27166
/rpt_family="AT_rich"
repeat_region 27737..27834
/rpt_family="MLT1H"
repeat_region 28328..28506
/rpt_family="MLT1K"
repeat_region 28513..28541
/rpt_family="AT_rich"
repeat_region 28813..28858
/rpt_family="AT_rich"
repeat_region 29627..30186
/rpt_family="L1_MN"
repeat_region 30465..30578
/rpt_family="T-rich"
repeat_region complement(30581..34457)
/rpt_family="L1_MN"
repeat_region complement(34455..35895)
/rpt_family="L1"
repeat_region complement(36134..36736)
/rpt_family="Lx4"
repeat_region 36749..36920
/rpt_family="Lx2"
repeat_region 37617..37751
/rpt_family="TA)n"
repeat_region 37958..37986
/rpt_family="(GAA)n"

```

Query Match 65.1%; Score 678; DB 10; Length 164759;
 Best Local Similarity 80.0%; Pred. No. 2.8e-170;
 Matches 824; Conservative 0; Mismatches 200; Indels 6; Gaps 2;

```

QY 12 TCGGCTAGCAGATGGAGTACTCGTCTCCACCTCCAGCTCGGTAAAGGTGGCTGT 71
Db 14611 TTGTAAGCAAGGAGGATGGCTGTCTGTCTAGCTGAGTAAGGTTGGCTAT 146052
QY 72 GCGACTCGGGCCATTTGGTGGAGCAACAGCGGAGCAAGTATCCCCCTGTGTGCGGG 131
Db 146051 ATAATTGGCTTTTGGTGGATGAGACAGACAGAGGAGGACCCCTGTGTACGAGC 145992
QY 132 CATGACAGCTGCTCTTAGAGATTGTCTAACTGGAGGAAACACAGGAGACTCTCAATA 191
Db 145991 CATAGACAGCTGCTCTTGAAGTGGCTAACTGGAGGAAATACAGGAGACTATCAATA 145932

```

```

QY 192 CCAGTTTTCATGCTTCTATGGGAGAGGAGTACTCAGCAGACATCTATCAGGTTTCAGT 251
Db 145931 TCAGTTTTCATGCTTCCATGGAGAGAGGACATCAGTGGGACATCTATGAGTTTGT 145872
QY 252 GCAGCCCATCTAAGGCACTTCTGGAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 311
Db 145871 ACAGCCCATCTAAGGCACTTCTGGAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 145812
QY 312 CACAGGAGCTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 371
Db 145811 GACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 145754
QY 372 GCGGCTCTCATGAGACCTCTGACAGTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 431
Db 145753 TCAGACTCTCCAGACCTCTGACAGTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 145694
QY 432 GCGGCTCTCTGACACCTCTGACAGTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 491
Db 145693 GGCATTTCTGTTGCTTTGCTATTAGATCTCTTAACAGGAGGAGGAGGAGGAGGAGGAG 145634
QY 492 GGCCTCTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 551
Db 145633 GGATCCTGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 145574
QY 552 GGGTCTCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 611
Db 145573 AGGCTCTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 145514
QY 612 CAGTCGAAATCGGAGTGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 671
Db 145513 CAGTAGAATCGAGTGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 145454
QY 672 TGTGCTCTGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 731
Db 145453 AGCGCTCTGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 145394
QY 732 AAACTCTACCTGATTTGAGTGGGTGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 787
Db 145393 AAACTCTACCTGATTTGAGTGGGTGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 145334
QY 788 AGGGCTTTGGCTAAAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 847
Db 145333 AGGGCTTTGGCTAAAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 145274
QY 848 TGGTAGATGCTGATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 907
Db 145273 TAGTGGATGCTAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 145214
QY 908 GCCTATTGCGAGGAGTCTCTGGGTGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 967
Db 145213 GTCTGTTGCAAGAGTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 145154
QY 968 CTGAGAGAGGCTTTTACCTAGACAGGAGTCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1027
Db 145153 CTGAAAGAGGATTTTACCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 145094
QY 1028 AGGTGATCAA 1037
Db 145093 AGGTGATCAA 145084

```

```

RESULT 14
AC101908/c
LOCUS
DEFINITION Mus musculus chromosome 1 clone RP24-275J1 map 1, *** SEQUENCING IN
PROGRESS ***, 8 unordered pieces.
AC101908
AC101908.4 GI:44199137
VERSION HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

```


REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

1 (bases 1 to 297639)
Birren,B., Nusbaum,C. and Lander,E.
Mus musculus chromosome 1, clone RP24-275J1
Unpublished
2 (bases 1 to 297639)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 297639)

REFERENCE
AUTHORS

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepe,Y., Collymore,A., Cooke,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.

TITLE
JOURNAL

Submitted (18-JUN-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 27, 2004 this sequence version replaced gi:31880232.

COMMENT

All repeats were identified using RepeatMasker:

Smith,A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: L17732

Center clone name: 275_J_1

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved

* 1 63747: contig of 63747 bp in length
* 63748 63847: gap of 100 bp
* 63848 73892: contig of 10045 bp in length
* 73893 73992: gap of 100 bp
* 73993 154450: contig of 80458 bp in length
* 154451 154550: gap of 100 bp
* 154551 167808: contig of 13258 bp in length
* 167809 167809: gap of 100 bp
* 167810 176711: contig of 8803 bp in length
* 176712 233500: contig of 56689 bp in length
* 233501 241222: contig of 7622 bp in length
* 241223 241322: gap of 100 bp
* 241323 297639: contig of 56317 bp in length.

FEATURES

source

1..297639
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/map="1"
/clone="RP24-275J1"
/clone_lib="RPCI-24 Male Mouse BAC"

ORIGIN

Query Match 62.5%; Score 650.8; DB 2; Length 297639;
Best Local Similarity 78.3%; Pred. No. 5.3e-163;
Matches 807; Conservative 0; Mismatches 217; Indels 6; Gaps 2;
QY 12 TCGGTAAGCAAGATTGGAGCTACTCGTCTCCACCTTCAGCTCGCGTAAGGTTGGCTGT 71
Db 1179 TTGTATAAGCAGAGTAGGATTCGGCTGCTCTGCTTCTAGCTGGAGTAAGGTTGGCTAT 1120
QY 72 GCGACTCGCGCCATTGTGTGATGACACGGGAGCAAGTGTATCCCTCTGTGCGGGG 131
Db 1119 ATATTTGTGCTGTTTGTGTGATGAGAGACAGAAAGCAAGGACACCCCTGTGTACGAGC 1060
QY 132 CATGGACAGCTGCTCTTAGAGATTCTTAACCTGGAGAACCCACAGAGACTCTCAATA 191
Db 1059 CATAGACAGCTGCTCTTGTGAAGTGTCTAACTGGAGGAATACAGAGACTATCAATA 1000
QY 192 CCAGTTTGTATGCTTCTTATGGGAGAGAGTACTACAGAGGACATCTATGAGTTTCAGT 251
Db 999 TCAGTTTGTATGCTTCCATGGAGAGAGACACTCAGTGGGACATCTATGTAGTTTAGT 940
QY 252 GCAGCCCATCTTAAGGCACTTCTGGAAGGCAAGATGCCAGTGTCTTGCCTATGGACC 311
Db 939 ACAGCCCATCTACTAAGGAACCTTGTGGAAGGCAAGATGCCAGTGTCTTGCCTATGGACC 880

312 CACAGAGCTGGGAAGACGACACAATGTCTGGGAGCCAGAGCAACCTGGGTTGATCCC 371
Db 879 GACTGGGAGGGAAGA--CACACATGTTGGGACGCCAGACAGCCTGGAGTATCC 822
QY 372 GCGGGCTCTCATGGACCTCTCTGAGCTCAAGAGGAGGAGGTGCCGAGGGCCGGCCATG 431
Db 821 TCAGACTCTCCAGACCTCTGCACTCAACAGGGAAGAGATGCTGAGGGCTGGCCATG 762
QY 432 GGGCCCTTTCTGTCACCATGTCTTACCTAGAGATCTACACAGAGAGAGATATTAGACCTCT 491
Db 761 GGACATTTCTGTGCTTTTGTGCTATTTAGATCTCTAACAGGAATATAGTATTAGACCTCTT 702
QY 492 GCAGCCTGCTTCGGGAGACCTGTGTAATCCGAGAAGACTCCGGGGGAATATCTGATTCC 551
Db 701 GGATCTCTTATAGTGTCACTTAATGTGACGGCCAGGACGAGGGAATACTGATTCC 642
QY 552 GGGTCTCTCCAGAGGCCCATCAGTAGCTTTGCTGATTTTGGCGGCACCTTCTCGCCAGC 611
Db 641 AGGCTCTCACAGAGAGCCCATCAGTAGTTTCTCTGACTTCGAGAGACACTTCTCTTAGT 582

```

QY 612 CAGTCGAATCGAGCTAGGAGCCACCGGCTCAACGAGCGCTCTCCCGAGTCATGC 671
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 581 CAGTGAATAATCAGTGTAGGAGCTACTCTGCTTAAACGAGCGCTCTTCCGAGTCATGC 522
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 672 TGTGCTCTCTGCTCAAGGTGGACCGAGGGAAGCTTTGGCCCCCATTTCCCGCAGCGAGGG 731
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 521 AGCGCTCTTGGTCAGAGTAGATCAGCATGAGCTTGTGACTCCATTTCCCGAGTGGGAAG 462
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 732 AAACTCTACCTGATGATCTTGGCTGGGTGAGAGCAACCC-----GGCGCAGGCAACA 787
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 461 AAACTCTACCTTATGATTTGGCTGGGTTCAGAGCAACCCGTCAGTCGCAAGCAACC 402
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 788 AGGCGCTTCGGCTAAAGAGAGTGAGGCATCAACACCTCCCTGTTCTGCTGGGCAAG 847
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 401 AGGCGATCTGCTCAAGAGAGTGAGGCATCAACACCTCCCTGTTCTGCTGGGCAAG 342
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 848 TGTGATAGTCGCTGAATCAGGCGCTCCCTGCTGTACTTATCGGACAGCAAGCTCACTC 907
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 341 TAGTGATGATTAACACCGGCTCCCTCATATACCATACAGGACAGGAGCTCACTC 282
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 908 GCCTATTGAGAGCTCTCTGCTGGGTGCTCAGCCACAGTATCTTATTCGCAACATTGCC 967
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 281 GTCCTTGCAGACTCTCTGAGAGGCTCAGCTCATAGCATCTCTATTGTGAGCATTTGCTC 222
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 968 CTGAGAGCGCTTCTACCTAGACAGCTCGGCACTCAACTGCTGCCAGGTCGAAGG 1027
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 221 CTGAAGATGTTTTCCTGCGGATCAATCTCGAGCATTAACATCTACTGCTAGGTCGAAGG 162
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1028 AGGTGATCAAA 1037
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 161 AGGTGATTA 152
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
AC102127
LOCUS AC102127 196674 bp DNA linear ROD 01-MAY-2004
DEFINITION Mus musculus chromosome 1, clone RP23-202A19, complete sequence.
ACCESSION AC102127.11 GI:46931440
VERSION AC102127.11
KEYWORDS HTG.
SOURCE Mus musculus
ORGANISM Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 196674)
AUTHORS Birren,B., Nusbaum,C. and Lander,B.
TITLE Mus musculus chromosome 1, clone RP23-202A19
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 196674)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepeil,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marguis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome

```

REFERENCE
AUTHORS

Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 196674)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepeil,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu.C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.

Direct Submission

Submitted (10-APR-2004) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS

4 (bases 1 to 196674)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepeil,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu.C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.

Direct Submission

Submitted (01-MAY-2004) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On May 1, 2004 this sequence version replaced gi:4638225.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WIRB

Web site: http://www-seq.wi.mit.edu

Contact: sequence submissions@broad.mit.edu

----- Project Information

Center project name: L18071

Center clone name: 202_A_19

FEATURES
source

Location/Qualifiers
1. 196674
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/map="1"
/clone_lib="RP23-202A19"
/clone="RP23-202A19" Female Mouse BAC"

TITLE
JOURNAL

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome

```
misc_feature 1. .4229
/note="wgs_end_extension
clone_end:SP6"
repeat_region 1. .193
/rpt_family="MTA"
unsure 614. .746
/note="single clone coverage"
701. .706
/note="<30 qual single clone coverage"
repeat_region 1208. .1291
/rpt_family="(CTA)n"
repeat_region 1571. .1842
/rpt_family="RLTRETN MM"
repeat_region complement(1974. .2067)
/rpt_family="B4"
repeat_region 3219. .3319
/rpt_family="GA-rich"
repeat_region 3368. .3448
/rpt_family="(CA)n"
repeat_region 3626. .3688
/rpt_family="(CAAA)n"
repeat_region 3925. .4374
/rpt_family="Lx"
misc_feature 4230. .4235
/note="clone boundary
clone_end:SP6
site:EcoRI"
repeat_region 4383. .4411
/rpt_family="AT-rich"
repeat_region 4445. .4496
/rpt_family="(TCTCTG)n"
repeat_region 4509. .4638
/rpt_family="(TG)n"
repeat_region 4741. .4784
/rpt_family="(TG)n"
repeat_region 4785. .4832
/rpt_family="(TA)n"
repeat_region 4833. .4955
/rpt_family="(TG)n"
repeat_region 5398. .5431
/rpt_family="(GAAA)n"
repeat_region complement(5434. .5561)
/rpt_family="RMER15"
repeat_region complement(6185. .6272)
/rpt_family="RSINE1"
repeat_region 6516. .6570
/rpt_family="(T)n"
repeat_region complement(6571. .6823)
/rpt_family="Li_MM"
repeat_region 6821. .6939
/rpt_family="Li_MM"
repeat_region 7007. .7285
/rpt_family="Lx5"
repeat_region 7292. .7848
/rpt_family="Lx5"
repeat_region 8187. .8286
/rpt_family="RSINE1"
repeat_region complement(8823. .9061)
/rpt_family="B4A"
repeat_region complement(9243. .9815)
/rpt_family="Li"
repeat_region 9832. .9861
/rpt_family="(A)n"
repeat_region 9862. .10368
/rpt_family="Lx2"
repeat_region complement(10373. .11559)
/rpt_family="Li_MM"
repeat_region 11555. .12196
/rpt_family="Lx2"
repeat_region complement(13221. .13381)
/rpt_family="B3A"
repeat_region 13437. .13578
/rpt_family="B1_MM"
```

```
repeat_region 15222. .15254
/rpt_family="AT-rich"
repeat_region complement(15315. .15605)
/rpt_family="RMER20"
repeat_region complement(16008. .16517)
/rpt_family="RMERS"
repeat_region 16673. .16713
/rpt_family="(TG)n"
repeat_region 16809. .16866
/rpt_family="(TG)n"
repeat_region complement(16922. .17052)
/rpt_family="B4"
repeat_region 17101. .17490
/rpt_family="MLT2B1"
repeat_region 17725. .17762
/rpt_family="(GGAA)n"

Query Match 50.2%; Score 522.8; DB 10; Length 196674;
Best Local Similarity 74.4%; Pred. No. 9.5e-129;
Matches 726; Conservative 0; Mismatches 242; Indels 8; Gaps 5;

QY 69 TGTGGAGCTGCGCCATTTTGTGGATGGAACAGCGGGGAGCAAGTGATCCCCCTGTCTGGC 128
DB 101632 TTTATGACTATGCCCATTTTATGGATGAAGAAAAAGGAGGATGCCCTTGTGTAAG 101691

QY 129 GGGCATGGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACCAACAGAGACTCTCAA 188
DB 101692 AGGCATATAAGCTGCTCTTTGAGGTGCCACCTTGAGGAATCTCTATGAGATATAA 101751

QY 189 ATACAGTTTGTATGCTTCTATGGGAGAGAGTAGTACTAGCAGGAGCATCTATGACAGTTT 248
DB 101752 ATATCAGTTTAAATGCTTTTATTAGAGAGAGACACTAGCAGAGAACATCTAAGTAGGTT 101811

QY 249 AGTGAGCCCCATCTAAGCACTTCTGGAAGGGAGAGATGCCAGTGTCTTGCCTATG 308
DB 101812 AGTACAGCTATCTCTGTGCACTTCTGTGGATGGGAGAGATAACAGTGTCTTACATATG 101871

QY 309 ACCACAGGAGCTGGGAAGACGACACACATGCTGGGCAGCCC---AGAGCAACCTGGGG 364
DB 101872 GACTATTGAGGCAGGGAGACACATATAATCTGCGGTAGCCCCATGGGAACACATGGAG 101931

QY 365 TGATCCCGGGCTCTCATGAGACCTCTGACACTCAAGGAGAGAGGTGCGAGGGCC 424
DB 101932 TGATTCCTTAGCTCTCATAGACCTCTGACAGTAACAGGGAAGAGTGTGTAAGGCT 101991

QY 425 GGCATGGCCCTTTCTGTCAACATGCTTACTAGAGATCTACAGGAGAGAGTATTAG 484
DB 101992 GGCCATGGGACATTTCTGTAGGATGCTCTATTAGAGGTCTACAGGAAGAGTATTAG 102051

QY 485 ACCTCTGGACCCCTGCTTGGGAGACCTGGTAATCCGAGAAGACTGCCGGGGAGATATCC 544
DB 102052 ACCTCTTATATCTGATCAGAGACCTAGTGTATGAGAGATGTGCGAGGGAACATCC 102111

QY 545 TGATTCGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTGATTTTGAAGCGCATTC 604
DB 102112 TGATTCAGTTCTCACCCAGGACCCCAATCTGCTTCTCTGACTTTGAGCAGTACTT-C 102170

QY 605 TGCCAGCCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGCA 664
DB 102171 TGCCAGCCTGTAGAAATCAAGTTGTAGGAGTACTCTGCTTAAAGCAGTCTCTTACCATA 102230

QY 665 GTCATGCTGTGCTCTGCTCAAGGTGGACCGGGAAGCTTTGGCCCATTTTGGCCAGC 724
DB 102231 -TCATTCATGCTCTTTTATCAAGTTGGATCAGATGGACATTTGACTCTGTTTCACTAGT 102289

QY 725 GAGAGGAAAACTCTACCTGATTTGCTGGTGGTCAAGAGCAACCGCGCACAGGCA 784
DB 102290 GAAAGGAAAACTCTACCATATTGATTTGGTGGTTCAGAGGACCAACAACA-AGGCA 102348

QY 785 ACAGGGCCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCTCTCTTGTCTCTGGGCA 844
DB 102349 ACCAGTGCAATTCAACTCAAGAGAGTGGGCCCATCAACACCT-GCTTCTTGTACTGGGCA 102407
```

Search completed: November 10, 2004, 13:18:05
Job time : 4604.34 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2004, 09:45:10 ; Search time 538.26 Seconds
(without alignments)
10152.424 Million cell updates/sec

Title: US-10-797-893-5
Perfect score: 1041
Sequence: 1 aagggtcggtcggtgaag.....ccaaggagtgatcaattga 1041

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1041	100.0	1041	10	ADC23341 DNA encod
2	1041	100.0	1041	12	ADG60231 Human mlc
3	1038.4	99.8	1464	10	ADC23343 DNA encod
4	1038.4	99.8	1464	12	ADG60233 Human mlc
5	1038	99.7	1115	10	ADC23337 DNA encod
6	1038	99.7	1115	12	ADG60227 Human mlc
7	1035.4	99.5	1538	10	ADC23339 DNA encod
8	1035.4	99.5	1538	12	ADG60229 Human mlc
9	1035.4	99.5	2097	12	ADQ09241 Human KNS
10	1035.4	99.5	2099	10	ADK40995
11	1034	99.3	2104	3	AAP15853
12	265	25.5	464	9	ACH44493 Human foe
13	233.2	22.4	531	12	ACH77408 Human gen
14	230.4	22.1	232	12	ACH91108 Human gen
15	184.6	17.7	2095	11	ADMO1564 Human cDN
16	184.6	17.7	3624	12	ADMO4181 DNA encod
17	155.8	15.0	3374	8	ADJ95073 Novel NOV
18	154.8	14.9	2375	8	ABX34596 Human mld
19	136.6	13.1	1152	6	ABA94616 Nucleotid
20	136.6	13.1	1152	6	ABA94616 DNA encod
21	136.6	13.1	4108	6	ABA94614 Human kin

22	136.6	13.1	4108	8	ABS57215
23	135.2	13.0	2675	10	ADJ95077
24	127.8	12.3	897	4	AAI70088
25	127.8	12.3	897	6	AAL13619
26	127.8	12.3	897	12	ADM81051
27	127.6	12.3	482	10	ABT40796
28	125.2	12.0	1026	6	AAQ73061
29	125.2	12.0	1026	9	AAI56806
30	125.2	12.0	1026	10	ADB66785
31	124.2	11.9	1011	6	ABQ73062
32	124.2	11.9	1011	9	AAL56807
33	124.2	11.9	1011	10	ADB66787
34	121.4	11.7	1014	6	ABA94615
35	118.2	11.4	2132	10	ADC30338
36	115	11.0	1839	10	ADC30569
37	114.6	11.0	2034	4	ABL16733
38	114.6	11.0	3661	6	AAJ28568
39	114.6	11.0	3694	10	ADC10189
40	114.6	11.0	4034	4	ABL16732
41	114	11.0	3570	12	ADQ19012
42	113.6	10.9	987	6	AAJ28569
43	110.8	10.6	4292	10	ADJ95085
44	109.6	10.5	2394	4	ABL28591
45	104	10.0	1391	5	AAH89871

ALIGNMENTS

RESULT 1
ADC23341
ID ADC23341 standard; DNA; 1041 BP.
XX
AC ADC23341;
XX
DT 18-DEC-2003 (first entry)
XX
DE DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 5).
XX
KW human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;
KW cytoskeletal; cardiac; immunomodulator; antiinflammatory; gene therapy;
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;
KW cardiac hypertrophy; immune disorder; inflammation.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS I..1041
FT /*tag= a
FT /product= "KID protein"
XX
PN US6387644-B1.
XX
PD 14-MAY-2002.
XX
PF 28-NOV-2000; 2000US-00724224.
XX
PR 20-APR-1999; 99US-00295612.
PR 20-JUN-2000; 2000US-00597292.
XX
(CYTO-) CYTOKINETICS INC.
PI Beraud C;
XX
DR WPI; 2003-706919/67.
DR P-PSDB; ADC23342.
XX
PT Identifying a candidate agent as modulator of function of a target
PT protein for treating cellular proliferation disorders by adding a
PT candidate agent to a mixture of the target protein that
PT directly/indirectly produces ADP or phosphate.
XX
PS Disclosure; SEQ ID NO 5; 26pp; English.

XX This invention relates to a novel method for high throughput screening
 CC systems used to identify compounds for the treatment of cellular
 CC proliferation disorders. Specifically, it refers to candidate agents that
 CC are capable of modulating the activity of target proteins having motor
 CC domains, such that the target protein directly or indirectly produces ADP
 CC or phosphate. Furthermore, this activity can be determined using
 CC fluorescence or absorbance readouts. The present invention describes a
 CC method that identifies modulators of the target protein, which is a
 CC kinesin-like DNA binding protein (known as KID) as cytoskeletal, cardants,
 CC immunomodulators and antiinflammatory. Accordingly, through gene
 CC therapy, they can be used for the treatment of cancer, hyperplasias,
 CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This
 CC polynucleotide sequence is human KID DNA (SeqID 5) encoding a full length
 CC KID enzyme of the invention.

XX Sequence 1041 BP; 228 A; 295 C; 302 G; 216 T; 0 U; 0 Other;

Query Match 100.0%; Score 1041; DB 10; Length 1041;
 Best Local Similarity 100.0%; Pred. No. 5e-296;
 Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGTGCTGTCGGCTAAGCAAGATGGAGCTACTCGTCTGCTCCACCTCCAGCTCGCGTA 60
 DB 1 ATGGGTGCTGTCGGCTAAGCAAGATGGAGCTACTCGTCTGCTCCACCTCCAGCTCGCGTA 60

QY 61 AGGGTGGCTGTCGGCTGCGGCTGTCGATGGAACGCGGAGCAAGTGATCCCCC 120
 DB 61 AGGGTGGCTGTCGGCTGCGGCTGTCGATGGAACGCGGAGCAAGTGATCCCCC 120

QY 121 TGTGTGCGGGGATGCGACAGCTGCTCTAGAGATGCTAACTGAGGAACCCACGAGG 180
 DB 121 TGTGTGCGGGGATGCGACAGCTGCTCTAGAGATGCTAACTGAGGAACCCACGAGG 180

QY 181 ACTCTCAAAATACAGTTTGTATGCTTCTATGCGGAGAGAGTACTCAGCAGACATCTAT 240
 DB 181 ACTCTCAAAATACAGTTTGTATGCTTCTATGCGGAGAGAGTACTCAGCAGACATCTAT 240

QY 241 GCAGGTTGAGTGCAGGCTTCTTAAGGCACTGCTGGAAGGCGCAAGTGCAGTGTGCTT 300
 DB 241 GCAGGTTGAGTGCAGGCTTCTTAAGGCACTGCTGGAAGGCGCAAGTGCAGTGTGCTT 300

QY 301 GCCTATGACCCACAGGAGCTGGGAAGAGCGACACAATGCTGGGAGCCAGAGCAACT 360
 DB 301 GCCTATGACCCACAGGAGCTGGGAAGAGCGACACAATGCTGGGAGCCAGAGCAACT 360

QY 361 GGGGTGATCCCGCGGCTCTCATGACCTCTGACAGTCTCAAGGGAGGAGGTCGCGAG 420
 DB 361 GGGGTGATCCCGCGGCTCTCATGACCTCTGACAGTCTCAAGGGAGGAGGTCGCGAG 420

QY 421 GGCGGGCCATGGGCGCTTCTGTACCATGCTTACCTAGAGATCTACGAGGAGGTA 480
 DB 421 GGCGGGCCATGGGCGCTTCTGTACCATGCTTACCTAGAGATCTACGAGGAGGTA 480

QY 481 TTAGACCTCTGAGCTCTGCGGAGACCTGGAATCCGAGAGAGCTGCGGGGGAAT 540
 DB 481 TTAGACCTCTGAGCTCTGCGGAGACCTGGAATCCGAGAGAGCTGCGGGGGAAT 540

QY 541 ATCTCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGTGATTTTTCAGCGGAC 600
 DB 541 ATCTCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGTGATTTTTCAGCGGAC 600

QY 601 TTCCTGCCAGCGAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACAGGCTCTCC 660
 DB 601 TTCCTGCCAGCGAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACAGGCTCTCC 660

QY 661 CGCAGTCTGCTGCTCTGCTCAAGGTGACACGAGGGAACGTTTGGCCCCATTTCCG 720
 DB 661 CGCAGTCTGCTGCTCTGCTCAAGGTGACACGAGGGAACGTTTGGCCCCATTTCCG 720

QY 721 CAGCGAGAGGGAATCTTACTGATGACTTTGGTGGGTGAGGACAACCGGCGACA 780
 DB 721 CAGCGAGAGGGAATCTTACTGATGACTTTGGTGGGTGAGGACAACCGGCGACA 780

QY 781 GGCAACAGGCGCTTCGGCTAAAGAGAGAGTGGAGCCATCAACACCTCCTGTTTCTCTG 840
 DB 781 GGCAACAGGCGCTTCGGCTAAAGAGAGAGTGGAGCCATCAACACCTCCTGTTTCTCTG 840

QY 841 GGCAAGTGGTAGATGCGCTGAATCAGGGCTCCCTCGTGTACCTTATCGGAGACAGCAAG 900
 DB 841 GGCAAGTGGTAGATGCGCTGAATCAGGGCTCCCTCGTGTACCTTATCGGAGACAGCAAG 900

QY 901 CTCACCTGCGCTTATTCAGGAGCTCTCTGCGTGGCTCAGCCACAGTATCCTTATTCGCAAC 960
 DB 901 CTCACCTGCGCTTATTCAGGAGCTCTCTGCGTGGCTCAGCCACAGTATCCTTATTCGCAAC 960

QY 961 ATTGCGCTCAGAGACGCTTCTACCTAGACACAGTCTCGGACTCAACTTTGCTGCCAGG 1020
 DB 961 ATTGCGCTCAGAGACGCTTCTACCTAGACACAGTCTCGGACTCAACTTTGCTGCCAGG 1020

QY 1021 TCCAGGAGGTGATCAATTGA 1041
 DB 1021 TCCAGGAGGTGATCAATTGA 1041

RESULT 2
 ADQ60231
 ID ADQ60231 standard; DNA; 1041 BP.
 XX AC ADQ60231;
 XX DT 23-SEP-2004 (first entry)
 XX Human microtubule motor protein DNA #3.
 XX Human; microtubule motor protein; gene; ds;
 KW cellular proliferation disorder; cancer; hyperplasia; restenosis;
 KW cardiac hypertrophy; immune disorder; inflammation;
 KW kinesin-like DNA binding protein; KID; autoimmune disease; arthritis;
 KW graft rejection; inflammatory bowel disease; angioplasty.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT 1.1041
 FT /tag= a
 FT /product= "Human microtubule motor protein #3"
 XX US6762043-B1.
 XX 13-JUL-2004.
 XX 06-MAR-2002; 2002US-00093317.
 XX 20-APR-1999; 99US-00295612.
 XX 20-JUN-2000; 2000US-00597292.
 XX 28-NOV-2000; 2000US-00724224.
 XX (CYTO-) CYTOKINETICS INC.
 XX Beraud C;
 XX WPI; 2004-532491/51.
 XX P-PSDB; ADQ60232.
 XX New isolated microtubule motor protein, useful for screening modulators
 PT for treating cellular proliferation disorders such as cancer,
 PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
 PT inflammation.
 XX Disclosure; SEQ ID NO 5; 26pp; English.
 XX The invention relates to human microtubule motor proteins and the nucleic
 CC acids encoding them. The invention also relates to a method of screening
 CC for modulators of a motor protein which has microtubule stimulated ATPase
 CC activity, a method of testing for ATPase activity of microtubule motor

CC proteins, methods to identify candidate agents that bind to a target
 CC protein or act as a modulator of the binding characteristics or
 CC biological activity of a target protein, modulators of the target
 CC protein, and methods of treating cellular proliferation disorders such as
 CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
 CC and inflammation, for treating disorders associated with kinesin-like DNA
 CC binding protein (KID) and for inhibiting KID. The sequences are used for
 CC screening for modulators of motor proteins useful for treating cellular
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
 CC hypertrophy, immune disorders and inflammation, for treating disorders
 CC associated with KID and for inhibiting KID and for treating autoimmune
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and
 CC proliferation induced after medical procedures including surgery and
 CC angioplasty. This sequence represents DNA encoding a human microtubule
 CC motor protein of the invention.
 XX
 SQ Sequence 1041 BP; 228 A; 295 C; 302 G; 216 T; 0 U; 0 Other;

Query Match 100.0%; Score 1041; DB 12; Length 1041;
 Best Local Similarity 100.0%; Pred. No. 5e-296;
 Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGTGCGTGTGGCTAAGCAAGATTGGAGTACTGCTGCTCCACCTCCAGCTCGGTA 60
 DB 1 ATGGGTGCGTGTGGCTAAGCAAGATTGGAGTACTGCTGCTCCACCTCCAGCTCGGTA 60
 QY 61 AGGGTGGCTGTGGGACTCGCGCCCATTTGTGGATGGAAACAGCGGGAGCAAGTATCCCCC 120
 DB 61 AGGGTGGCTGTGGGACTCGCGCCCATTTGTGGATGGAAACAGCGGGAGCAAGTATCCCCC 120
 QY 121 TGTGTGGGGGATGAGACAGTCTCTCTAGAGATTGCTTAATCGGAGAACACACAGGAG 180
 DB 121 TGTGTGGGGGATGAGACAGTCTCTCTAGAGATTGCTTAATCGGAGAACACACAGGAG 180
 QY 181 ACTCTCAATACAGTTGATGCTCTTATGGGAGAGGAGTACTCAGCAGGACATCTAT 240
 DB 181 ACTCTCAATACAGTTGATGCTCTTATGGGAGAGGAGTACTCAGCAGGACATCTAT 240
 QY 241 GCAGGTTTCAGTCAGCGCCCATCTTAAGCACTTCTGGAAGGGCAGAGTGCAGTGTGCTT 300
 DB 241 GCAGGTTTCAGTCAGCGCCCATCTTAAGCACTTCTGGAAGGGCAGAGTGCAGTGTGCTT 300
 QY 301 GCCTATGGACCCACAGAGTGGGAAGAGCACAACAATGCTGGCGAGCCCGAGCAACCT 360
 DB 301 GCCTATGGACCCACAGAGTGGGAAGAGCACAACAATGCTGGCGAGCCCGAGCAACCT 360
 QY 361 GGGGTGATCCCGGGCTCTCATGGACCTCTCTGAGCTCACAAGGAGGAGGTGCGAG 420
 DB 361 GGGGTGATCCCGGGCTCTCATGGACCTCTCTGAGCTCACAAGGAGGAGGTGCGAG 420
 QY 421 GCGCGGCCATGGGCCCTTTCTGTACCATGCTTACTAGAGATCTACAGGAGAGAGTA 480
 DB 421 GCGCGGCCATGGGCCCTTTCTGTACCATGCTTACTAGAGATCTACAGGAGAGAGTA 480
 QY 481 TTAGACCTCTGGAACCTGCTTCGGAGACCTGGTAATCCGAGAAGACTGCCGGGGAAT 540
 DB 481 TTAGACCTCTGGAACCTGCTTCGGAGACCTGGTAATCCGAGAAGACTGCCGGGGAAT 540
 QY 541 ATCTGATTCGGGCTCTCCAGAGACCCATCAGTAGCTTGTGATTTAGCGGCAC 600
 DB 541 ATCTGATTCGGGCTCTCCAGAGACCCATCAGTAGCTTGTGATTTAGCGGCAC 600
 QY 601 TTCCTCCAGCAGTGGAAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCC 660
 DB 601 TTCCTCCAGCAGTGGAAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCC 660
 QY 661 CGCAGTCATGCTGTGCTCTGCTGCAAGTGGAGCCAGCGGAACTTTGGCCCCCATTTCCG 720
 DB 661 CGCAGTCATGCTGTGCTCTGCTGCAAGTGGAGCCAGCGGAACTTTGGCCCCCATTTCCG 720
 QY 721 CAGCGAGAGGGAACCTCTACTGATTTGATCTGGGTGAGTCAAGACACCGCGGCACA 780
 DB 721 CAGCGAGAGGGAACCTCTACTGATTTGATCTGGGTGAGTCAAGACACCGCGGCACA 780

QY 781 GGCAACAAGGGCTTGGCTAAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTG 840
 DB 781 GGCAACAAGGGCTTGGCTAAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTG 840
 QY 841 GGCAAGTGTAGATGCGCTGAATCAAGGGCTCCCTGCTAGCTTATCGGAGACAGCAAG 900
 DB 841 GGCAAGTGTAGATGCGCTGAATCAAGGGCTCCCTGCTAGCTTATCGGAGACAGCAAG 900
 QY 901 CTCACCTGCGCTATTGAGGAGTCTCTGGGTGGCTCAGCCACAGTATCCTTATTTGCCAAC 960
 DB 901 CTCACCTGCGCTATTGAGGAGTCTCTGGGTGGCTCAGCCACAGTATCCTTATTTGCCAAC 960
 QY 961 ATTGCCCTCGAGAGAGCTTCTTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGG 1020
 DB 961 ATTGCCCTCGAGAGAGCTTCTTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGG 1020
 QY 1021 TCCAAGGAGTGAATCAATTGA 1041
 DB 1021 TCCAAGGAGTGAATCAATTGA 1041

RESULT 3

ADC23343
 ID ADC23343 standard; DNA; 1464 BP.
 XX
 AC ADC23343;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 7).
 XX
 KW human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;
 KW cytoskeletal; cardiant; immunomodulator; antiinflammatory; gene therapy;
 KW cancer; hyperplasia; restenosis; cellular proliferation disorder;
 KW cardiac hypertrophy; immune disorder; inflammation.
 XX
 OS Homo sapiens.
 XX
 Key Location/Qualifiers
 CDS 1..1464
 FT /*tag= a
 FT /product= "KID protein"
 XX
 PN US6387644-B1.
 XX
 PD 14-MAY-2002.
 XX
 PF 28-NOV-2000; 2000US-00724224.
 XX
 PR 20-APR-1999; 99US-00295612.
 PR 20-JUN-2000; 2000US-00597292.
 XX
 PA (CYTO-) CYTOKINETICS INC.
 XX
 PI Beraud C;
 XX
 DR WPI; 2003-706919/67.
 DR P-PSDB; ADC23344.
 XX

PT Identifying a candidate agent as modulator of function of a target
 PT protein for treating cellular proliferation disorders by adding a
 PT candidate agent to a mixture of the target protein that
 PT directly/indirectly produces ADP or phosphate.
 XX
 PS Disclosure; SEQ ID NO 7; 26pp; English.
 XX
 CC This invention relates to a novel method for high throughput screening
 CC systems used to identify compounds for the treatment of cellular
 CC proliferation disorders. Specifically, it refers to candidate agents that
 CC are capable of modulating the activity of target proteins having motor
 CC domains, such that the target protein directly or indirectly produces ADP
 CC or phosphate. Furthermore, this activity can be determined using

CC fluorescence or absorbance readouts. The present invention describes a
 CC method that identifies modulators of the target protein, which is a
 CC kinesin-like DNA binding protein (known as KID) as cytosolic, cardinals,
 CC immunomodulators and antiinflammatory. Accordingly, through gene
 CC therapy, they can be used for the treatment of cancer, hyperplasia,
 CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This
 CC polynucleotide sequence is human KID DNA (SeqID 7) encoding a full length
 CC KID enzyme of the invention.
 XX
 SQ Sequence 1464 BP; 342 A; 409 C; 428 G; 285 T; 0 U; 0 Other;

Query Match 99.8%; Score 1038.4; DB 10; Length 1464;
 Best Local Similarity 99.9%; Pred. No. 3.4e-295;
 Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGTCTGCTGGCTAAGCAAGATTGGAGTACTCTGCTGCTCACTCCAGCTCGCGTA 60
 DB 1-ATGGGTCTGCTGGCTAAGCAAGATTGGAGTACTCTGCTGCTCACTCCAGCTCGCGTA 60

QY 61 AGGGTGGCTGTGCGACTCGGCGCATTTCTGATGGAACAGCGGGAGCAAGTATCCGCC 120
 DB 61 AGGGTGGCTGTGCGACTCGGCGCATTTCTGATGGAACAGCGGGAGCAAGTATCCGCC 120

QY 121 TGTGTGCGGGGATGAGCAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCCAGGAG 180
 DB 121 TGTGTGCGGGGATGAGCAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCCAGGAG 180

QY 181 ACTCTCAAAATACCAATTTGATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTAT 240
 DB 181 ACTCTCAAAATACCAATTTGATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTAT 240

QY 241 GCAGTTCAGTGCACGCCATCTTAGGCACCTTCTGGAAGGCGCAATGCCAGTGTGCTT 300
 DB 241 GCAGTTCAGTGCACGCCATCTTAGGCACCTTCTGGAAGGCGCAATGCCAGTGTGCTT 300

QY 301 GCCTATGACCCACAGGAGCTGGGAAGAGCGACACAAATGCTGGGAGCCCGAGCAACT 360
 DB 301 GCCTATGACCCACAGGAGCTGGGAAGAGCGACACAAATGCTGGGAGCCCGAGCAACT 360

QY 361 GGGGTGATCCCGGGGCTCTATGACCTCTGAGACCTCTGAGCTCACAGGAGAGGGTCCGAG 420
 DB 361 GGGGTGATCCCGGGGCTCTATGACCTCTGAGACCTCTGAGCTCACAGGAGAGGGTCCGAG 420

QY 421 GSCCGGCATCGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTA 480
 DB 421 GSCCGGCATCGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTA 480

QY 481 TTAGACCTCTGGAACCTGCTTGGGAGACCTGGTAATCCGAGAGACTGCGGGGGAAT 540
 DB 481 TTAGACCTCTGGAACCTGCTTGGGAGACCTGGTAATCCGAGAGACTGCGGGGGAAT 540

QY 541 ATCTGTATCCGGGTCTCTCCAGAGCGCCATCAGTAGCTTGTGATTTGAGCGGCAC 600
 DB 541 ATCTGTATCCGGGTCTCTCCAGAGCGCCATCAGTAGCTTGTGATTTGAGCGGCAC 600

QY 601 TTCTGCGCAGCAGCTGGAATCGGACTGTAGAGACCCCGCTCAACAGCGCTCTCC 660
 DB 601 TTCTGCGCAGCAGCTGGAATCGGACTGTAGAGACCCCGCTCAACAGCGCTCTCC 660

QY 661 CGCAGTCTATGCTGCTCTCTGCTCAAGGTGGAACAGCGGGAACGTTTGGCCCCATTCCG 720
 DB 661 CGCAGTCTATGCTGCTCTCTGCTCAAGGTGGAACAGCGGGAACGTTTGGCCCCATTCCG 720

QY 721 CAGCAGAGGGAATCTTACTGATGATCTGCTGGTCTGAGAGACACCGGCGACA 780
 DB 721 CAGCAGAGGGAATCTTACTGATGATCTGCTGGTCTGAGAGACACCGGCGACA 780

QY 781 GGCACAAAGGGCTTCGGGTAAGAGAGTGGAGCATCAACACCTCTGTTTCTCTG 840
 DB 781 GGCACAAAGGGCTTCGGGTAAGAGAGTGGAGCATCAACACCTCTGTTTCTCTG 840

QY 841 GGCACAAAGTGTAGTGCCTGTAATCAGGGCCTCTCTGCTGATACCTTATCGGAGACGAG 900
 DB 841 GGCACAAAGTGTAGTGCCTGTAATCAGGGCCTCTCTGCTGATACCTTATCGGAGACGAG 900

Db 841 GGCACAAAGTGTAGTGCCTGTAATCAGGGCCTCTCTGCTGATACCTTATCGGAGACGAG 900
 QY 901 CTCACTCGCTATTGTCAGGACTCTCTGGTGGCTTCAGCCACAGTATCTTATTGCCAAC 960
 Db 901 CTCACTCGCTATTGTCAGGACTCTCTGGTGGCTTCAGCCACAGTATCTTATTGCCAAC 960
 QY 961 ATTGCCCCCTGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAG 1020
 Db 961 ATTGCCCCCTGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAG 1020

QY 1021 TCCAAGGAGGTGATCAATTG 1040
 Db 1021 TCCAAGGAGGTGATCAATCG 1040

RESULT 4
 ADQ60233
 ID ADQ60233 standard; DNA; 1464 BP.
 AC ADQ60233;
 XX
 XX 23-SEP-2004 (first entry)
 DT
 XX Human microtubule motor protein DNA #4.
 DE
 XX Human; microtubule motor protein; gene; ds;
 KW cellular proliferation disorder; cancer; hyperplasia; restenosis;
 KW cardiac hypertrophy; immune disorder; inflammation;
 KW kinesin-like DNA binding protein; KID; autoimmune disease; arthritis;
 KW graft rejection; inflammatory bowel disease; angioplasty.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1464
 FT /*tag= a
 FT /product= "Human microtubule motor protein #4"
 XX
 PN US6762043-B1.
 XX
 PD 13-JUL-2004.
 XX
 XX 06-MAR-2002; 2002US-00093317.
 XX
 XX 20-APR-1999; 99US-00295612.
 PR 20-JUN-2000; 2000US-00597292.
 PR 28-NOV-2000; 2000US-00724224.
 XX
 XX (CYTO-) CYTOKINETICS INC.
 XX
 XX Beraud C;
 XX
 XX WPI; 2004-532491/51.
 DR P-PSDB; ADQ60234.
 XX
 PT New isolated microtubule motor protein, useful for screening modulators
 PT for treating cellular proliferation disorders such as cancer,
 PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
 PT inflammation.
 XX
 XX Disclosure; SEQ ID NO 7; 26pp; English.
 XX
 CC The invention relates to human microtubule motor proteins and the nucleic
 CC acids encoding them. The invention also relates to a method of screening
 CC for modulators of a motor protein which has microtubule stimulated ATPase
 CC activity, a method of testing for ATPase activity of microtubule motor
 CC proteins, methods to identify candidate agents that bind to a target
 CC protein or act as a modulator of the binding characteristics of
 CC biological activity of a target protein, modulators of the target
 CC protein, and methods of treating cellular proliferation disorders such as
 CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
 CC and inflammation, for treating disorders associated with kinesin-like DNA
 CC binding protein (KID) and for inhibiting KID. The sequences are used for

CC screening for modulators of motor proteins useful for treating cellular
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
 CC hypertrophy, immune disorders and inflammation, for treating disorders
 CC associated with KID and for inhibiting KID and for treating autoimmune
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and
 CC proliferation induced after medical procedures including surgery and
 CC angioplasty. This sequence represents DNA encoding a human microtubule
 CC motor protein of the invention.

XX SQ Sequence 1464 BP; 342 A; 409 C; 428 G; 285 T; 0 U; 0 Other;

Query Match 99.8%; Score 1038.4; DB 12; Length 1464;
 Best Local Similarity 99.9%; Pred. No. 3.4e-295;
 Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCAGCTCCGGTA 60
 DB 1 ATGGGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCAGCTCCGGTA 60
 QY 61 AGGGTGGCTGTCGCACTGCGGCCATTGTTGGATGGAAACAGCGGGAGCAAGTGATCCCCC 120
 DB 61 AGGGTGGCTGTCGCACTGCGGCCATTGTTGGATGGAAACAGCGGGAGCAAGTGATCCCCC 120
 QY 121 TGTGTCGGGGCATGACAGCTGCTCTCTAGAGATTGCTTAACCTGGAGGAACCAACAGGAG 180
 DB 121 TGTGTCGGGGCATGACAGCTGCTCTCTAGAGATTGCTTAACCTGGAGGAACCAACAGGAG 180
 QY 181 ACTCTCAATACCACTTTGATGCTTCTATGGGGAGAGGAGTACTCAGCAGGACATCTAT 240
 DB 181 ACTCTCAATACCACTTTGATGCTTCTATGGGGAGAGGAGTACTCAGCAGGACATCTAT 240
 QY 241 GCAGGTTCACTAGTCAGGCCCATCTTCAAGGCCATCTGCTGGAAGGGCAGAAATGCCAGTGTCTT 300
 DB 241 GCAGGTTCACTAGTCAGGCCCATCTTCAAGGCCATCTGCTGGAAGGGCAGAAATGCCAGTGTCTT 300
 QY 301 GCTATGAGCCCAACAGAGCTGGGAAGAGCAACAAATGCTGGGCGAGCCCAAGCAACCT 360
 DB 301 GCTATGAGCCCAACAGAGCTGGGAAGAGCAACAAATGCTGGGCGAGCCCAAGCAACCT 360
 QY 361 GGGGTGATCCCGGGGCTCTCATGGACCTCTCATGGACCTCTCATGGACCTCTCATGGAGGAGGGTCCGAG 420
 DB 361 GGGGTGATCCCGGGGCTCTCATGGACCTCTCATGGACCTCTCATGGAGGAGGGTCCGAG 420
 QY 421 GCGCGGCCATGGGCCCTTTCTGTACCATGCTTTACCTAGAGATCTTACAGAGAGAGGTA 480
 DB 421 GCGCGGCCATGGGCCCTTTCTGTACCATGCTTTACCTAGAGATCTTACAGAGAGAGGTA 480
 QY 481 TTAGACTCTGTGACCTCTGCTGGGAGACCTGGTAATTCGAGAGACTGCGGGGGAAT 540
 DB 481 TTAGACTCTGTGACCTCTGCTGGGAGACCTGGTAATTCGAGAGACTGCGGGGGAAT 540
 QY 541 ATCTCTGATTCGGGCTCTTCCAGAGAGCCCATCAGTAGCTTTGCTGATTTTGAGCGGCAC 600
 DB 541 ATCTCTGATTCGGGCTCTTCCAGAGAGCCCATCAGTAGCTTTGCTGATTTTGAGCGGCAC 600
 QY 601 TTCCTGCCAGCCAGTCCGAATCGEATGTAGAGCCACCGGCTCAACAGCGCTCCTCC 660
 DB 601 TTCCTGCCAGCCAGTCCGAATCGEATGTAGAGCCACCGGCTCAACAGCGCTCCTCC 660
 QY 661 CGGAGTCATGCTGCTCTGCTCAAGGTGGACAGCGGGAACGTTTGGCCCCCATTTCCG 720
 DB 661 CGGAGTCATGCTGCTCTGCTCAAGGTGGACAGCGGGAACGTTTGGCCCCCATTTCCG 720
 QY 721 CAGCGAGAGGGAACCTCTACCTGATTGCTGGCTGGGTCAGAGCAACCGGGCCACA 780
 DB 721 CAGCGAGAGGGAACCTCTACCTGATTGCTGGCTGGGTCAGAGCAACCGGGCCACA 780
 QY 781 GGCACAAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTTTCCTG 840
 DB 781 GGCACAAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTTTCCTG 840
 QY 841 GGCACAAGTGGTAGTCGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGACAGCAAG 900

DB 841 GGCACAAGTGGTAGTCGCTGAATCAGGGCCCTCCCTGCTAGCTTATCGGACAGCAAG 900
 QY 901 CTCACCTCGCTATTGAGGACTCTCTGGTGGCTCAGCCACAGTAGTATCTTATGCAAC 960
 DB 901 CTCACCTCGCTATTGAGGACTCTCTGGTGGCTCAGCCACAGTAGTATCTTATGCAAC 960
 QY 961 ATTGCCCTCGAGAGAGCGTTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGG 1020
 DB 961 ATTGCCCTCGAGAGAGCGTTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGG 1020
 QY 1021 TCCAGGAGGTGATCAATTG 1040
 DB 1021 TCCAGGAGGTGATCAATCG 1040

RESULT 5

ADC23337 standard; DNA; 1115 BP.

XX ADC23337;

AC AC (first entry)

DT 18-DEC-2003

XX DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 1).
 DE human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;
 KW cytosolic; cardiant; immunomodulator; antiinflammatory; gene therapy;
 KW cancer; hyperplasia; restenosis; cellular proliferation disorder;
 KW cardiac hypertrophy; immune disorder; inflammation.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1115

FT /*tag= a

FT /product= "KID protein"

FT /transl_except= (pos: 1..5; aa: Met)

FT /note= "This codon has an apparent 2 nucleotide insertion
 that alters the reading frame"

XX US6397644-B1.

XX 14-MAY-2002.

XX 28-NOV-2000; 2000US-00724224.

XX 20-APR-1999; 99US-00295612.

XX 20-JUN-2000; 2000US-00597292.

XX (CYTO-) CYTOKINETICS INC.

XX Beraud C;

XX WPI; 2003-706919/67.

XX P-PSDB; ADC23338.

XX Identifying a candidate agent as modulator of function of a target
 XX protein for treating cellular proliferation disorders by adding a
 XX candidate agent to a mixture of the target protein that
 XX directly/indirectly produces ADP or phosphate.

XX Disclosure; SEQ ID NO 1; 26pp; English.

XX This invention relates to a novel method for high throughput screening
 XX systems used to identify compounds for the treatment of cellular
 XX proliferation disorders. Specifically, it refers to candidate agents that
 XX are capable of modulating the activity of target proteins having motor
 XX domains, such that the target protein directly or indirectly produces ADP
 XX or phosphate. Furthermore, this activity can be determined using
 XX fluorescence or absorbance readouts. The present invention describes a
 XX method that identifies modulators of the target protein, which is a
 XX kinesin-like DNA binding protein (known as KID) as cytosolic, cardants,
 XX immunomodulators and antiinflammatories. Accordingly, through gene

CC therapy, they can be used for the treatment of cancer, hyperplasias, CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This CC polynucleotide sequence is human KID DNA (Seqid 1) encoding a full length CC KID enzyme of the invention.

XX	Sequence	1115 BP; 241 A; 318 C; 333 G; 223 T; 0 U; 0 Other;	
XX	Query Match	99.7%; Score 1038; DB 10; Length 1115;	
XX	Best Local Similarity	100.0%; Pred. No. 3.9e-295;	
XX	Matches 1038; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	4	GGTGGCTGTCGCTAAGCAAGATTGGAGCTACTGCTGTCACCTCCAGCTCGCGTAAGG	63
DB	78	GGTGGCTGTCGCTAAGCAAGATTGGAGCTACTGCTGTCACCTCCAGCTCGCGTAAGG	137
QY	64	GTGGCTGTGGCACTGCGGCCATTTGTGATGGAACAGCGGAGCAAGTGTATCCCTCTGT	123
DB	138	GTGGCTGTGGCACTGCGGCCATTTGTGATGGAACAGCGGAGCAAGTGTATCCCTCTGT	197
QY	124	GTGGGGGATGGAGCAGCTGCTCTTAGAGATTGCTAACTGGAGGAACCAACGAGACT	183
DB	198	GTGGGGGATGGAGCAGCTGCTCTTAGAGATTGCTAACTGGAGGAACCAACGAGACT	257
QY	184	CTCAATACCAAGTTGATGCTCTCTATGGGAGGAGTACTCAGCAGCATCTATGCA	243
DB	258	CTCAATACCAAGTTGATGCTCTCTATGGGAGGAGTACTCAGCAGCATCTATGCA	317
QY	244	GTTTCAGTGCAGCCCATCTAAGGCATTTGCTGGAAGGGCAGAAATGCCAGTGTCTGCC	303
DB	318	GTTTCAGTGCAGCCCATCTAAGGCATTTGCTGGAAGGGCAGAAATGCCAGTGTCTGCC	377
QY	304	TATGAGCCACAGAGAGCTGGGAAGACGACACAAATGCTGGGAGCCGAGACCACTGGG	363
DB	378	TATGAGCCACAGAGAGCTGGGAAGACGACACAAATGCTGGGAGCCGAGACCACTGGG	437
QY	364	GTGATCCCGGGCTCTCATGACCTCTGACAGCTCACAAGGGAGAGGGTGGCGAGGGC	423
DB	438	GTGATCCCGGGCTCTCATGACCTCTGACAGCTCACAAGGGAGAGGGTGGCGAGGGC	497
QY	424	CGGCCATGGGCTTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGAGGTATTA	483
DB	498	CGGCCATGGGCTTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGAGGTATTA	557
QY	484	GACCTCTGGACCTGCTTCGGGAGACCTGGTAATCCGAGAGACTGCGGGGAAATATC	543
DB	558	GACCTCTGGACCTGCTTCGGGAGACCTGGTAATCCGAGAGACTGCGGGGAAATATC	617
QY	544	CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGTATTTGAGCGGCACTTC	603
DB	618	CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGTATTTGAGCGGCACTTC	677
QY	604	CTGCCAGCAGTCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAAT	663
DB	678	CTGCCAGCAGTCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAAT	737
QY	664	AGTCATGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	723
DB	738	AGTCATGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	797
QY	724	CGAGAGGGAACCTCTACCTGATTTGATTTGGTGGGTTCAGAGCAACCCGGGCGCAGGC	783
DB	798	CGAGAGGGAACCTCTACCTGATTTGATTTGGTGGGTTCAGAGCAACCCGGGCGCAGGC	857
QY	784	AACAGGGCTTGGCTAAAGAGTGGAGCCATCAACCTCTCTGTTTCTCTGGGC	843
DB	858	AACAGGGCTTGGCTAAAGAGTGGAGCCATCAACCTCTCTGTTTCTCTGGGC	917
QY	844	AAAGTGTAGTGCCTGATACAGGGCTCCCTCTGCTGATCTATCGGACGACGAGCTC	903
DB	918	AAAGTGTAGTGCCTGATACAGGGCTCCCTCTGCTGATCTATCGGACGACGAGCTC	977
QY	904	ACTGGCTATTGAGGAGCTCTCTGGGTGGCTGAGCCACAGATATCTTTATTCACACAT	963

Db	978	ACTCGCTATTGCGAGCTCTCTGGGTGGCTAGCCACAGTATCTTATTGCCAACAAT	1037
QY	964	GCCCTTGAGAGAGCGTTCTACTAGACAGACTCTCGGACTCACTTCTGTCGAGTCC	1023
Db	1038	GCCCTTGAGAGAGCGTTCTACTAGACAGACTCTCGGACTCACTTCTGTCGAGTCC	1097
QY	1024	AAGGAGGTGATCAATGA	1041
Db	1098	AAGGAGGTGATCAATGA	1115
RESULT 6			
ID	ADQ60227	standard; DNA; 1115 BP.	
XX	AC	ADQ60227;	
XX	AC	23-SEP-2004 (first entry)	
XX	AC	Human microtubule motor protein DNA #1.	
XX	DE	Human; microtubule motor protein; gene; ds;	
XX	KW	cellular proliferation disorder; cancer; hyperplasia; restenosis;	
XX	KW	cardiac hypertrophy; immune disorder; inflammation;	
XX	KW	kinesin-like DNA binding protein; KID; autoimmune disease; arthritis;	
XX	KW	graft rejection; inflammatory bowel disease; angioplasty.	
XX	OS	Homo sapiens.	
XX	PN	US6762043-B1.	
XX	PD	13-JUL-2004.	
XX	PF	06-MAR-2002; 2002US-00093317.	
XX	PR	20-APR-1999; 99US-00295612.	
XX	PR	20-JUN-2000; 2000US-00597292.	
XX	PR	28-NOV-2000; 2000US-00724224.	
XX	PA	(CYTO-) CYTOKINETICS INC.	
XX	PI	Beraud C;	
XX	XX	WPI; 2004-532491/51.	
XX	DR	New isolated microtubule motor protein, useful for screening modulators	
XX	PT	for treating cellular proliferation disorders such as cancer,	
XX	PT	hyperplasias, restenosis, cardiac hypertrophy, immune disorders and	
XX	PT	inflammation.	
XX	PS	Disclosure; SEQ ID NO 1; 26pp; English.	
XX	CC	The invention relates to human microtubule motor proteins and the nucleic	
XX	CC	acids encoding them. The invention also relates to a method of screening	
XX	CC	for modulators of a motor protein which has microtubule stimulated ATPase	
XX	CC	activity, a method of testing for ATPase activity of microtubule motor	
XX	CC	proteins, methods to identify candidate agents that bind to a target	
XX	CC	protein or act as a modulator of the binding characteristics or	
XX	CC	biological activity of a target protein, modulators of the target	
XX	CC	protein, and methods of treating cellular proliferation disorders such as	
XX	CC	cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders	
XX	CC	and inflammation, for treating disorders associated with kinesin-like DNA	
XX	CC	binding protein (KID) and for inhibiting KID. The sequences are used for	
XX	CC	screening for modulators of motor proteins useful for treating cellular	
XX	CC	proliferation disorders such as cancer, hyperplasias, restenosis, cardiac	
XX	CC	hypertrophy, immune disorders and inflammation, for treating disorders	
XX	CC	associated with KID and for inhibiting KID and for treating autoimmune	
XX	CC	diseases, arthritis, graft rejection, inflammatory bowel disease and	
XX	CC	proliferation induced after medical procedures including surgery and	
XX	CC	angioplasty. This sequence represents DNA encoding a human microtubule	
XX	CC	motor protein of the invention. Note: The specification states that this	
XX	CC	sequence encodes the protein featured as SEQ ID NO:2, but this does not	
XX	CC	appear to be the case.	

XX	Sequence	1115 BP; 241 A; 318 C; 333 G; 223 T; 0 U; 0 Other;	
SQ	Query Match	99.7%; Score 1038; DB 12; Length 1115;	
	Best Local Similarity	100.0%; Pred. No. 3.9e-295;	
	Matches 1038; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Qy	4	GTCGCTGTCGCTAAGCAAGATTGGAGTACTCGTCCACCTCCAGCTCGCGTAAGG	63
Db	78	GGTCGCTGTCGCTAAGCAAGATTGGAGTACTCGTCCACCTCCAGCTCGCGTAAGG	137
Qy	64	GTGGCTGTGGACTCGCGCCATTGTTGGATGGAACAGCGGAGCAAGTATCCCTCGT	123
Db	138	GTGGCTGTGGACTCGCGCCATTGTTGGATGGAACAGCGGAGCAAGTATCCCTCGT	197
Qy	124	GTGGGGGCAATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACACCCAGGAGCT	183
Db	198	GTGGGGGCAATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACACCCAGGAGCT	257
Qy	184	CTCAATACCAAGTTTGAATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	243
Db	258	CTCAATACCAAGTTTGAATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	317
Qy	244	GTTTCAGTCAGCCCATCTTAAGCAGCTTCTGGAAGGCGAGATGCCAGTGTGCTTGGC	303
Db	318	GGTTTCAGTCAGCCCATCTTAAGCAGCTTCTGGAAGGCGAGATGCCAGTGTGCTTGGC	377
Qy	304	TATGGACCCACAGGAGCTGGGAGAGCAGACACAAATGCTGGCGACCCAGAGCAACCTGGG	363
Db	378	TATGGACCCACAGGAGCTGGGAGAGCAGACACAAATGCTGGCGACCCAGAGCAACCTGGG	437
Qy	364	GTGATCCCGCGGCTCTCATGGACCTCTCGAGCTCAAGGGAGGAGGTCGCGAGGCG	423
Db	438	GTGATCCCGCGGCTCTCATGGACCTCTCGAGCTCAAGGGAGGAGGTCGCGAGGCG	497
Qy	424	CGGCCATGGGCGCTTTCTGTACCATGTCTTACCTAGAGATCTACGAGGAGAGGTATTA	483
Db	498	CGGCCATGGGCGCTTTCTGTACCATGTCTTACCTAGAGATCTACGAGGAGAGGTATTA	557
Qy	484	GACCTCTCGACCTGCTTCGGGAGACTGCTTAATCCGAGAAGACTGCGGGGGAATATC	543
Db	558	GACCTCTCGACCTGCTTCGGGAGACTGCTTAATCCGAGAAGACTGCGGGGGAATATC	617
Qy	544	CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGAAGCGGCACCTC	603
Db	618	CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGAAGCGGCACCTC	677
Qy	604	CTGCCAGCCAGTCGAATCGGACTGTAGAGCCACCGGCTCAACAGCGCTCTCTCCCGC	663
Db	678	CTGCCAGCCAGTCGAATCGGACTGTAGAGCCACCGGCTCAACAGCGCTCTCTCCCGC	737
Qy	664	AGTCATGCTGCTCTCTGCTCAAGGTGGACCGGGGACGTTTGGCCCATTTGCCAG	723
Db	738	AGTCATGCTGCTCTCTGCTCAAGGTGGACCGGGGACGTTTGGCCCATTTGCCAG	797
Qy	724	CGAGAGGAAACTCTACCTGATTGACTTTGGCTGGGTGAGAGCAACCGGCGCACAGGC	783
Db	798	CGAGAGGAAACTCTACCTGATTGACTTTGGCTGGGTGAGAGCAACCGGCGCACAGGC	857
Qy	784	AACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCTCCCTGTTTGTCTGGGC	843
Db	858	AACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCTCCCTGTTTGTCTGGGC	917
Qy	844	AAAGTGTAGATCGCGTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGACACAGAGTC	903
Db	918	AAAGTGTAGATCGCGTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGACACAGAGTC	977
Qy	904	ACTCGCTATTGCGAGCTCTCTGGGTGGCTCAGCCACAGATATCTTATTGCAACATT	963
Db	978	ACTCGCTATTGCGAGCTCTCTGGGTGGCTCAGCCACAGATATCTTATTGCAACATT	1037
Qy	964	GCCCTGTAGAGAGCGCTTCTACCTAGACAGAGTCTCCGACTCAACTTGTGTCGAGGTCC	1023

Query Match

99.5%; Score 1035.4; DB 10; Length 1538;

Db	1038	GCCCTCGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGGTCC	1097
Qy	1024	AAGGAGGTGATCAATTGA	1041
Db	1098	AAGGAGGTGATCAATTGA	1115
RESULT 7			
ADC23339			
ID	ADC23339	standard; DNA; 1538 BP.	
XX	AC	ADC23339;	
XX	DT	18-DEC-2003 (first entry)	
XX	DE	DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 3).	
XX	KW	human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;	
KW	KW	cytostatic; cardiac; immunomodulator; antiinflammatory; gene therapy;	
KW	KW	cancer; hyperplasia; restenosis; cellular proliferation disorder;	
KW	KW	cardiac hypertrophy; immune disorder; inflammation.	
OS	OS	Homo sapiens.	
XX	Key	Location/Qualifiers	
PH	FT	1..1538	
FT	FT	/*tag= a	
FT	FT	/product= "KID protein"	
FT	FT	/transl_except= (pos: 4..5; aa: Pro)	
FT	FT	/note= "This codon has an apparent 1 nucleotide deletion	
FT	FT	that alters the reading frame"	
XX	XX	US6387644-B1.	
XX	XX	14-MAY-2002.	
XX	XX	28-NOV-2000; 2000US-00724224.	
XX	XX	20-APR-1999; 99US-00295612.	
XX	XX	20-JUN-2000; 2000US-00597292.	
XX	XX	(CYTO-) CYTOKINETICS INC.	
XX	XX	Beraud C;	
XX	XX	WPI; 2003-706919/67.	
XX	XX	P-PSDB; ADC23340.	
XX	XX	Identifying a candidate agent as modulator of function of a target	
XX	XX	protein for treating cellular proliferation disorders by adding a	
XX	XX	candidate agent to a mixture of the target protein that	
XX	XX	directly/indirectly produces ADP or phosphate.	
XX	XX	Disclosure; SEQ ID NO 3; 26pp; English.	
XX	XX	This invention relates to a novel method for high throughput screening	
XX	XX	systems used to identify compounds for the treatment of cellular	
XX	XX	proliferation disorders. Specifically, it refers to candidate agents that	
XX	XX	are capable of modulating the activity of target proteins having motor	
XX	XX	domains, such that the target protein directly or indirectly produces ADP	
XX	XX	or phosphate. Furthermore, this activity can be determined using	
XX	XX	fluorescence or absorbance readouts. The present invention describes a	
XX	XX	method that identifies modulators of the target protein, which is a	
XX	XX	kinesin-like DNA binding protein (known as KID) as cytostatic, cardiac,	
XX	XX	immunomodulators and antiinflammatories. Accordingly, through gene	
XX	XX	therapy, they can be used for the treatment of cancer, hyperplasias,	
XX	XX	restenosis, cardiac hypertrophy, immune disorders and inflammation. This	
XX	XX	polynucleotide sequence is human KID DNA (SeqID 3) encoding a full length	
XX	XX	KID enzyme of the invention.	
XX	XX	Sequence 1538 BP; 355 A; 432 C; 459 G; 292 T; 0 U; 0 Other;	

Best Local Similarity 99.9%; Pred. No. 2.6e-294;		Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	4	GGTCGCTGTCGGCTAAGCAAGATTGGAGTACTCGTCGTCCACCTCCAGCTCCGCTAAGG	63
Db	78	GGTCGCTGTCGGCTAAGCAAGATTGGAGTACTCGTCGTCCACCTCCAGCTCCGCTAAGG	137
QY	64	GTGGCTGTGCGACTGCGGCCATTGTGTGATGGAACAGCGGGAGCAAGTGATCCCCCTGT	123
Db	138	GTGGCTGTGCGACTGCGGCCATTGTGTGATGGAACAGCGGGAGCAAGTGATCCCCCTGT	197
QY	124	GTGCGGGGATGAGCAGCTGCTCTTAGAGATTGCTAAGGAGCAACACAGGAGACT	183
Db	198	GTGCGGGGATGAGCAGCTGCTCTTAGAGATTGCTAAGGAGCAACACAGGAGACT	257
QY	184	CTCAATACCACTTTGATGCCCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA	243
Db	258	CTCAATACCACTTTGATGCCCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA	317
QY	244	GGTTCAGTGCAGCCATCTAAGGCACTTGTGTGGAAGGCGAGATGCCAGTGTGCTTGC	303
Db	318	GGTTCAGTGCAGCCATCTAAGGCACTTGTGTGGAAGGCGAGATGCCAGTGTGCTTGC	377
QY	304	TATGACCCACAGGAGCTGGGAGAGCGACACAATGCTGGGCGAGCCAGAGCAACTGG	363
Db	378	TATGACCCACAGGAGCTGGGAGAGCGACACAATGCTGGGCGAGCCAGAGCAACTGG	437
QY	364	GTGATCCCGGGCTCTCATGAGACCTCTGAGCTCACAAGGGAGGAGGTGCCAGGGC	423
Db	438	GTGATCCCGGGCTCTCATGAGACCTCTGAGCTCACAAGGGAGGAGGTGCCAGGGC	497
QY	424	CGGCCATGGGCCCTTCTGTACCATGTCTTACCTAGAGATCTACGAGGAGGTATTA	483
Db	498	CGGCCATGGGCCCTTCTGTACCATGTCTTACCTAGAGATCTACGAGGAGGTATTA	557
QY	484	GACCTCTGAGCCCTGCTCCGGAGACCTGTAATCCGAGAAGACTGCCGGGGATATC	543
Db	558	GACCTCTGAGCCCTGCTCCGGAGACCTGTAATCCGAGAAGACTGCCGGGGATATC	617
QY	544	CTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTC	603
Db	618	CTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTC	677
QY	604	CTGCCAGCAGTCCGAATCGACTGTAGAGCCACCGGCTCAACGAGCGCTCCCTCCGC	663
Db	678	CTGCCAGCAGTCCGAATCGACTGTAGAGCCACCGGCTCAACGAGCGCTCCCTCCGC	737
QY	664	AGTCATGCTGTCTCTGCTCAAGTGGACCGGGGAACTTTGGCCCTCATTTGCCAG	723
Db	738	AGTCATGCTGTCTCTGCTCAAGTGGACCGGGGAACTTTGGCCCTCATTTGCCAG	797
QY	724	CGAGAGGGAATCTTACCTGATTTGACTTGGCTGGGTGAGAGCAACCGGGCGCAGGC	783
Db	798	CGAGAGGGAATCTTACCTGATTTGACTTGGCTGGGTGAGAGCAACCGGGCGCAGGC	857
QY	784	AACAAGGCTTCGGCTAAGAGACTGGAGCCATCAACACCTCCCTGTTGCTCGGC	843
Db	858	AACAAGGCTTCGGCTAAGAGACTGGAGCCATCAACACCTCCCTGTTGCTCGGC	917
QY	844	AAAGTGTAGATCGCTGAATCAGGGCTCCCTCTGTGTATCTTATCGGACAGCAAGCTC	903
Db	918	AAAGTGTAGATCGCTGAATCAGGGCTCCCTCTGTGTATCTTATCGGACAGCAAGCTC	977
QY	904	ACTCGGCTATTGAGGACTCTCTGGGTGGCTAGCCCAAGTATCTTATGCAACATTT	963
Db	978	ACTCGGCTATTGAGGACTCTCTGGGTGGCTAGCCCAAGTATCTTATGCAACATTT	1037
QY	964	GCCCTGTAGAGCGCTTCTACCTAGACACAGTCTCCGACTCAACTTGTGTCAGGCTC	1023
Db	1038	GCCCTGTAGAGCGCTTCTACCTAGACACAGTCTCCGACTCAACTTGTGTCAGGCTC	1097
QY	1024	AAGGAGGTGATCAATTG 1040	

Db 1098 AAGGAGGTGATCAATCG 1114

RESULT 8
ADQ60229

ID ADQ60229 standard; DNA; 1538 BP.

XX ADQ60229;

AC ADQ60229;

XX 23-SEP-2004 (first entry)

XX Human microtubule motor protein DNA #2.

XX Human; microtubule motor protein; gene; ds;
cellular proliferation disorder; cancer; hyperplasia; restenosis;
cardiac hypertrophy; immune disorder; inflammation;
kinesin-like DNA binding protein; KID; autoimmune disease; arthritis;
graft rejection; inflammatory bowel disease; angioplasty.

OS Homo sapiens.

XX US6762043-B1.

XX 13-JUL-2004.

XX 06-MAR-2002; 2002US-00093317.

XX 20-APR-1999; 99US-00295612.

XX 20-JUN-2000; 2000US-00597292.

XX 28-NOV-2000; 2000US-00724224.

XX (CYTO-) CYTOKINETICS INC.

XX Beraud C;

XX WPI; 2004-532491/51.

XX New isolated microtubule motor protein, useful for screening modulators
for treating cellular proliferation disorders such as cancer,
hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
inflammation.

XX Disclosure; SEQ ID NO 3; 26pp; English.

XX The invention relates to human microtubule motor proteins and the nucleic
acids encoding them. The invention also relates to a method of screening
for modulators of a motor protein which has microtubule stimulated ATPase
activity, a method of testing for ATPase activity of microtubule motor
proteins, methods to identify candidate agents that bind to a target
protein or act as a modulator of the binding characteristics or
biological activity of a target protein, modulators of the target
protein, and methods of treating cellular proliferation disorders such as
cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
and inflammation, for treating disorders associated with kinesin-like DNA
binding protein (KID) and for inhibiting KID. The sequences are useful for
screening for modulators of motor proteins useful for treating cellular
proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
hypertrophy, immune disorders and inflammation, for treating disorders
associated with KID and for inhibiting KID and for treating autoimmune
diseases, arthritis, graft rejection, inflammatory bowel disease and
proliferation induced after medical procedures including surgery and
angioplasty. This sequence represents DNA encoding a human microtubule
motor protein of the invention. Note: The specification states that this
sequence encodes the protein featured as SEQ ID NO:4, but this does not
appear to be the case.

XX Sequence 1538 BP; 355 A; 432 C; 459 G; 292 T; 0 U; 0 Other;

Query Match 99.5%; Score 1035.4; DB 12; Length 1538;

Best Local Similarity 99.9%; Pred. No. 2.6e-294;

Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGTGCTGTGCGCTAAGCAAGATTGGAGTACTCGTCGTCACCTCCAGCTCCGCTAAGG 63

Db 78 GGTGCTGTCTGGCTTAAGCAGATTGGAGCTACTGCTGCTCAGCTCCAGCTCGCTAAGG 137
 QY 64 GTGGCTGTGGCACTGGCGCCATTGTGGATGGAAACAGCGGAGCAAGTATATCCCTGT 123
 Db 138 GTGGCTGTGGCACTGGCGCCATTGTGGATGGAAACAGCGGAGCAAGTATATCCCTGT 197
 QY 124 GTGGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACACACAGGAGACT 183
 Db 198 GTGGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACACACAGGAGACT 257
 QY 184 CTCAATACCACTTGTATGCTCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 243
 Db 258 CTCAATACCACTTGTATGCTCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 317
 QY 244 GGTTCAGTGCAGCCCATCTTAAGGCACTTCTGCGAGGGGAGAAATGCGAGTGCTTGGC 303
 Db 318 GGTTCAGTGCAGCCCATCTTAAGGCACTTCTGCGAGGGGAGAAATGCGAGTGCTTGGC 377
 QY 304 TATGGACCCACAGGAGCTGGGAAGACGACACAAATGCTGGGACGCCAGAGCAACCTGGG 363
 Db 378 TATGGACCCACAGGAGCTGGGAAGACGACACAAATGCTGGGACGCCAGAGCAACCTGGG 437
 QY 364 GTGATCCCGGGGCTCTCATGGACCTCTGCACTCTCAAGGGAGGAGGTGCGAGGGC 423
 Db 438 GTGATCCCGGGGCTCTCATGGACCTCTGCACTCTCAAGGGAGGAGGTGCGAGGGC 497
 QY 424 CGGCCATGGGCCCTTTCTGTACCATGCTTACTAGAGATCTACCAAGAGAGGTATTATTA 483
 Db 498 CGGCCATGGGCCCTTTCTGTACCATGCTTACTAGAGATCTACCAAGAGAGGTATTATTA 557
 QY 484 GACCTCTGAGACCTGCTTGGGAGACCTGGTAAATCCGAGAGACTGCCGGGGGAATATC 543
 Db 558 GACCTCTGAGACCTGCTTGGGAGACCTGGTAAATCCGAGAGACTGCCGGGGGAATATC 617
 QY 544 CTGATTCGGGCTCTCCGAAAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTC 603
 Db 618 CTGATTCGGGCTCTCCGAAAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTC 677
 QY 604 CTGCGAGCGAGTGAATCGGACTGTAGGAGCCACCCGCTCAACAGCGCTCTCCCGC 663
 Db 678 CTGCGAGCGAGTGAATCGGACTGTAGGAGCCACCCGCTCAACAGCGCTCTCCCGC 737
 QY 664 AGTCATGCTGTCTCTGCTCAAGGTGGACAGCGGGAACGTTTGGCCCAATTTGCCAG 723
 Db 738 AGTCATGCTGTCTCTGCTCAAGGTGGACAGCGGGAACGTTTGGCCCAATTTGCCAG 797
 QY 724 CGAGAGGAAACTCTTACCTGATTTGACTTGGGTGAGAGCAACCGGGGCACAGGC 783
 Db 798 CGAGAGGAAACTCTTACCTGATTTGACTTGGGTGAGAGCAACCGGGGCACAGGC 857
 QY 784 AACAGGGCTTGGCTTAAGAGAGTGGAGCCATCAACCTCCCTGTTGCTTGGGC 843
 Db 858 AACAGGGCTTGGCTTAAGAGAGTGGAGCCATCAACCTCCCTGTTGCTTGGGC 917
 QY 844 AAGTGTAGATGCTGATAGGCGCTCCCTCGGTGATCTTATCGGACAGCAAGCTC 903
 Db 918 AAGTGTAGATGCTGATAGGCGCTCCCTCGGTGATCTTATCGGACAGCAAGCTC 977
 QY 904 ACTGGCTATTGAGGACTCTCTGGGTGCTCAGGCCACAGTATCTTATGCGCAACTT 963
 Db 978 ACTGGCTATTGAGGACTCTCTGGGTGCTCAGGCCACAGTATCTTATGCGCAACTT 1037
 QY 964 GCCCTGAGAGCGCTTCTACCTAGACAGTCTCGGACTCAACTTCTGCGAGGTCC 1023
 Db 1038 GCCCTGAGAGCGCTTCTACCTAGACAGTCTCGGACTCAACTTCTGCGAGGTCC 1097
 QY 1024 AAGGAGTGCATCAATTG 1040
 Db 1098 AAGGAGTGCATCAATCG 1114

ADQ09241

ID ADQ09241 standard; cDNA; 2097 BP.

XX AC ADQ09241;

XX DT 23-SBP-2004 (first entry)

XX XX Human KNSL4 encoding cDNA SEQ ID NO:426.

XX DE thanatos-associated protein; THAP; THAP responsive gene; THAP family;
 XX KW THAP responsive element; angiogenesis; inflammation; metastasis; cancer;
 KW apoptosis; cardiovascular disease; neurodegenerative disease; chemokine;
 KW antiangiogenic; antiinflammatory; cardiovascular; cytostatic;
 KW neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;
 KW human; gene; ss.

XX OS Homo sapiens.

XX PN WO2004055050-A2.

XX PD 01-JUL-2004.

XX PF 10-DEC-2003; 2003WO-IB006434.

XX PR 10-DEC-2002; 2002US-0432699P.

XX PR 03-JUL-2003; 2003US-0485027P.

XX PA (ENDO-) ENDOCUBE SAS.

XX PA (CNRS) CNRS CENT NAT RECH SCI.

XX PI Girard J, Amalric F, Roussigne M, Clouaire T;

XX DR WPI: 2004-525034/50.

XX DR P-PSDB; ADQ09240.

XX PT Modulating expression of a Thanatos (death)-Associated Protein (THAP)
 responsive gene for preventing or treating e.g. cancer or inflammation,
 comprises modulating the interaction of a THAP polypeptide with a nucleic
 acid.
 XX PS Example 47; SEQ ID NO 426; 612pp; English.
 XX CC The present invention describes a method for modulating the expression of
 a thanatos (death)-associated protein (THAP) responsive gene. The method
 comprises modulating the interaction of a THAP-family polypeptide or its
 biological fragment with a nucleic acid, and so enhancing or repressing
 the expression of the THAP responsive gene. Also described: (1) a method
 of modulating the expression of a gene responsive to a THAP/chemokine
 complex; (2) a pharmaceutical composition comprising a THAP responsive
 element in a pharmaceutical carrier; (3) a transcription factor decoy
 consisting essentially of a THAP responsive element; (4) a cell
 comprising a transcription factor decoy described above; (5) methods of
 modulating the interaction between a nucleic acid and a THAP-family
 polypeptide or its biological fragment, or a nucleic acid and a
 THAP/chemokine complex; (6) a vector packaging cell line comprising a
 cell comprising a viral vector which comprises a promoter operably linked
 to a nucleic acid encoding a THAP-family polypeptide or its biological
 fragment; (7) a method of constructing a cell which expresses a
 recombinant THAP-family polypeptide; (8) a method of ameliorating
 symptoms associated with a condition mediated by a THAP/chemokine complex
 ; (9) methods of identifying a test compound that modulates transcription
 at a THAP responsive element or that modulates the transport of a
 chemokine into the nucleus; (10) methods for reducing the symptoms
 associated with a condition selected from excessive or insufficient
 angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
 or insufficient apoptosis, cardiovascular disease and neurodegenerative
 diseases; symptoms associated with a condition resulting from the
 activity of a chemokine or a THAP-family polypeptide in an individual; or
 symptoms associated with transcriptional repression or activation
 mediated by a THAP-family polypeptide in an individual; (11) a vector
 comprising a THAP responsive promoter operably linked to a nucleic acid
 encoding a detectable product; (12) a genetically engineered cell
 comprising the vector described above or that expresses a THAP-family

CC terminal tail; or (e) is a complement of (d). The nucleic acid molecules, CC polypeptides, methods and substance are useful for treating cancers, CC immune-related diseases or disorders, cardiovascular disease, brain or CC neuronal-associated diseases, and metabolic disorders. The disorders are CC preferably cancers of the tissues or of hematopoietic origin, diseases of CC the central or peripheral nervous system, Alzheimer's disease, CC Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis, CC viral infections, infections caused by prions, infections caused by CC bacteria, infections caused by fungi, ocular diseases, migraines, pain, CC sexual dysfunction, mood disorders, attention disorders, cognition CC disorders, hypotension, hypertension, psychotic disorders, neurological CC disorders, dyskinesias, metabolic disorders and organ transplant CC rejection. This sequence corresponds to the DNA encoding one of the CC kinase polypeptides of the invention.

XX Sequence 2099 BP; 484 A; 592 C; 628 G; 395 T; 0 U; 0 Other;

Query Match 99.5%; Score 1035.4; DB 10; Length 2099;
Best Local Similarity 99.9%; Pred. No. 3e-294;
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	4	GGTGGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG	63
DB	101	GGTGGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG	160
QY	64	GTGGCTGTCGCACTGCGGCCATTGTTGGATGGAACACGCGGAGCAAGTATCCCCCTGT	123
DB	161	GTGGCTGTCGCACTGCGGCCATTGTTGGATGGAACACGCGGAGCAAGTATCCCCCTGT	220
QY	124	GTGGGGGATGAGACAGCTGCTCTTAGAGATTGCTAAGTGGAGAACACACAGGAGACT	183
DB	221	GTGGGGGATGAGACAGCTGCTCTTAGAGATTGCTAAGTGGAGAACACACAGGAGACT	280
QY	184	CTCAATACACAGTTTTCATGCTTCTATGGGAGAGGAGTACTCAGCAGACATCTATGCA	243
DB	281	CTCAATACACAGTTTTCATGCTTCTATGGGAGAGGAGTACTCAGCAGACATCTATGCA	340
QY	244	GGTTCACTGTCAGCCCATCTTAAGGCATTTGCTGGAGGGCAGAAATGCCAGTGTGCTTGC	303
DB	341	GGTTCACTGTCAGCCCATCTTAAGGCATTTGCTGGAGGGCAGAAATGCCAGTGTGCTTGC	400
QY	304	TATGACCCACAGGAGCTGGAGAGCGACACATGCTGGGAGCCCGCAGAGCAACTGGG	363
DB	401	TATGACCCACAGGAGCTGGAGAGCGACACATGCTGGGAGCCCGCAGAGCAACTGGG	460
QY	364	GTGATCCCGCGGCTCTCATGACCTCTGAGCTCACAAGGAGGAGGAGGTCGCCAGGSC	423
DB	461	GTGATCCCGCGGCTCTCATGACCTCTGAGCTCACAAGGAGGAGGAGGTCGCCAGGSC	520
QY	424	CGGCCATGGGCGCTTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGAGGTATTA	483
DB	521	CGGCCATGGGCGCTTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGAGGTATTA	580
QY	484	GACCTCTGGAGCCCTGCTTGGGAGAGCTGGTAATCCGAGAGACTGCCGGGGGAAATATC	543
DB	581	GACCTCTGGAGCCCTGCTTGGGAGAGCTGGTAATCCGAGAGACTGCCGGGGGAAATATC	640
QY	544	CTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGCGCACTTC	603
DB	641	CTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGCGCACTTC	700
QY	604	CTGCCAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCTCCCGC	663
DB	701	CTGCCAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCTCCCGC	760
QY	664	AGTCATGCTGCTCTGCTGAAGTGGACAGCGGGAACGTTTGGCCCCCATTTGCCAG	723
DB	761	AGTCATGCTGCTCTGCTGAAGTGGACAGCGGGAACGTTTGGCCCCCATTTGCCAG	820
QY	724	CCAGAGGGGAAACTTACTGATTCAGTTGCTGCTGAGAGCAACCGGGGCAAGGC	783
DB	821	CCAGAGGGGAAACTTACTGATTCAGTTGCTGCTGAGAGCAACCGGGGCAAGGC	880

QY	784	AACAAGGGCTTCGGCTAAAAGAGAGTGGAGGCATCAACACCTCTCTTTTCTCTGGGC	843
DB	881	AACAAGGGCTTCGGCTAAAAGAGAGTGGAGGCATCAACACCTCTCTTTTCTCTGGGC	940
QY	844	AAAGTGGTAGATCGCTGAATCAGGGCCCTCCCTGCTGTAACCTTATCGGACAGCAAGCTC	903
DB	941	AAAGTGGTAGATCGCTGAATCAGGGCCCTCCCTGCTGTAACCTTATCGGACAGCAAGCTC	1000
QY	904	ACTCGCTATTTCAGGACTCTCTGGGTGGCTCAGGCCACAGTATCTTTATGCAACATT	963
DB	1001	ACTCGCTATTTCAGGACTCTCTGGGTGGCTCAGGCCACAGTATCTTTATGCAACATT	1060
QY	964	GCCCTCGAGAGAGCTTTTACCTAGACACAGTCTCCGCACTCAACTTTGTCGAGGTCC	1023
DB	1061	GCCCTCGAGAGAGCTTTTACCTAGACACAGTCTCCGCACTCAACTTTGTCGAGGTCC	1120
QY	1024	AAGGAGGTGATCAATTG	1040
DB	1121	AAGGAGGTGATCAATTG	1137

RESULT 11

AAFI5853	ID	AAFI5853 standard; cDNA; 2104 BP.
XX	AC	AAFI5853;
XX	AC	AC
XX	DT	13-MAR-2001 (first entry)
XX	DE	Human prostate cancer antigen nucleotide sequence SEQ ID NO:288.
XX	DE	Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW	KW	neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;
KW	KW	vulnary; gastrointestinal; nephrotropic; anti-infective; gynaecological;
KW	KW	antibacterial; gene therapy; neural; immune; reproductive; renal;
KW	KW	gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW	KW	wound; infectious disease; ss.
XX	OS	Homo sapiens.
XX	PN	WO200055174-A1.
XX	PD	21-SEP-2000.
XX	PF	08-MAR-2000; 2000WO-US005988.
XX	PR	12-MAR-1999; 99US-0124270P.
XX	PA	(HUMA-) HUMAN GENOME SCI INC.
XX	PA	(ROSE/) ROSEN C A.
XX	PI	Rosen CA, Ruben SM;
XX	DR	WPI; 2000-587513/55.
XX	DR	P-PSDB; AAB56650.
PT	PT	Prostate cancer associated gene sequences, referred to as prostate cancer
PT	PT	antigens, useful for treatment, prevention, and diagnosis of disorders
XX	XX	such as prostate cancer.
PS	PS	Claim 1; Page 805-806; 2338pp; English.
XX	CC	AAFI5566 to AAFI5605 encode the human prostate cancer associated
CC	CC	proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC	CC	The prostate cancer antigens can have neuroprotective, cytosolic,
CC	CC	cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
CC	CC	nephrotropic, anti-infective, gynaecological and antibacterial activities,
CC	CC	and can be used in gene therapy. The prostate cancer antigen
CC	CC	polynucleotides may be used for detection of prostate cancer, chromosome
CC	CC	identification, as chromosome markers, and for numerous other diagnostic
CC	CC	or research purposes. The prostate cancer antigens may be used to treat
CC	CC	disorders such as neural, immune, muscular, reproductive,
CC	CC	gastrointestinal, pulmonary, cardiovascular, renal, and proliferative

CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to	
CC AAB57303 represent sequences used in the exemplification of the present	
CC invention	
XX	
SQ Sequence 2104 BP; 517 A; 580 C; 597 G; 400 T; 0 U; 10 Other;	
Query Match 99.3%; Score 1034; DB 3; Length 2104;	
Best Local Similarity 99.7%; Pred. No. 7.7e-294;	
Matches 1034; Conservative 1; Mismatches 2; Indels 0; Gaps 0;	
QY	4 GGTGCTGTCGGCTAAGCAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGGTAAGG 63
DB	
QY	67 GGTGCTGTCGGCTAAGCAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGGTAAGG 126
DB	
QY	64 GTGGCTGTGCGACTGCGGCCAATTTGTGATGGAACAGCGGAGCAAGTGATGCCCTGT 123
DB	
QY	127 GTGGCTGTGCGACTGCGGCCAATTTGTGATGGAACAGCGGAGCAAGTGATGCCCTGT 186
DB	
QY	124 GTGGGGGCGATGAGCAGCTGCTCTAGAGATTGCTAACTGGAGAACCCACGAGACT 183
DB	
QY	187 GTGGGGGCGATGAGCAGCTGCTCTAGAGATTGCTAACTGGAGAACCCACGAGACT 246
DB	
QY	184 CTCGAATACCAAGTTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 243
DB	
QY	247 CTCGAATACCAAGTTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 306
DB	
QY	244 GGTTCAGTGCAGGCCATCCTAAGGCATTTGTGGAAGGGCAGAAATGCCAGTGTCTTGCC 303
DB	
QY	307 GGTTCAGTGCAGGCCATCCTAAGGCATTTGTGGAAGGGCAGAAATGCCAGTGTCTTGCC 366
DB	
QY	304 TATGACCCACAGGAGCTGGGAGAGCAGACAAATGCTGGCAGCCAGACCACTTGGG 363
DB	
QY	367 TATGACCCACAGGAGCTGGGAGAGCAGACAAATGCTGGCAGCCAGACCACTTGGG 426
DB	
QY	364 GTGATCCCGCGGCTCTCATGACCTCTCTGAGCTCACAAGGAGGAGGGTGCAGAGGC 423
DB	
QY	427 GTGATCCCGCGGCTCTCATGACCTCTCTGAGCTCACAAGGAGGAGGGTGCAGAGGC 486
DB	
QY	424 CGGCAATGGGCCCTTCTGTACCAATGCTTACCTAGAGATCTACCAGGAGAGGTATTA 483
DB	
QY	487 CGGCAATGGGCCCTTCTGTACCAATGCTTACCTAGAGATCTACCAGGAGAGGTATTA 546
DB	
QY	484 GACCTCTGGACCTCTCTGGGAGACCTGTGTAATCCGAGAAGCTGCGGGGGAATATC 543
DB	
QY	547 GACCTCTGGACCTCTCTGGGAGACCTGTGTAATCCGAGAAGCTGCGGGGGAATATC 606
DB	
QY	544 CTGATTCCGGGCTCTCTCCAGAAAGCCCATCAGTAGCTTTGTGATTTTGGCGGCATTC 603
DB	
QY	607 CTGATTCCGGGCTCTCTCCAGAAAGCCCATCAGTAGCTTTGTGATTTTGGCGGCATTC 666
DB	
QY	604 CTGCGAGCCAGTCCGAATCCGACTGTAGAGCCACCGGCTCAACGAGGCTCTCTCCGCG 663
DB	
QY	667 CTGCGAGCCAGTCCGAATCCGACTGTAGAGCCACCGGCTCAACGAGGCTCTCTCCGCG 726
DB	
QY	664 AGTCATGCTGTCTCTCTGCTCAAGGTGAGCAGCGGGAACGTTTGGCCCCCAATTTGCCAG 723
DB	
QY	727 AGTCATGCTGTCTCTCTGCTCAAGGTGAGCAGCGGGAACGTTTGGCCCCCAATTTGCCAG 786
DB	
QY	724 CGAGAGGGAAAACTCTACTGATTTGACTTTGGCTGGGTTCAGAGCAACCGGGGCGACAGC 783
DB	
QY	787 CGAGAGGGAAAACTCTACTGATTTGACTTTGGCTGGGTTCAGAGCAACCGGGGCGACAGC 846
DB	
QY	784 AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCTCTGTTTGTCTCTGGGC 843
DB	
QY	847 AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCTCTGTTTGTCTCTGGGC 906
DB	
QY	844 AAGTGGTAGATCGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGAGACGCAAGCTC 903
DB	
QY	907 AAGTGGTAGATCGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGAGACGCAAGCTC 966
DB	
QY	904 ACTCGCTATTGAGAGCTCTCTGGTGGCTCAGCCAGATCTCTTATTCGCAACATT 963
DB	
QY	967 ACTCGCTATTGAGAGCTCTCTGGTGGCTCAGCCAGATCTCTTATTCGCAACATT 1026

QY	964 GCGCTGAGAGACGCTTCTACCTAGACACAGTCTCGCAGCTCAACTTGTCTGCCAGGTCC 1023
DB	1027 GCGCTGAGAGACGCTTCTACCTAGACACAGTCTCGCAGCTCAACTTGTCTGCCAGGTCC 1086
QY	1024 AAGGAGGTGATCAATTG 1040
DB	1087 AAGGAGGTGATCAATCG 1103
RESULT 12	
ACH44493	
ID	ACH44493 standard; cDNA; 464 BP.
XX	
AC	ACH44493;
XX	
DT	13-OCT-2003 (first entry)
XX	
DE	Human foetal brain cDNA #5218.
XX	
KW	Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW	genome mapping; biodiversity; genetic disorder.
XX	
OS	Homo sapiens.
XX	
PN	US2003073623-A1.
XX	
PD	17-APR-2003.
XX	
PF	30-JUL-2001; 2001US-00918995.
XX	
PR	30-JUL-2001; 2001US-00918995.
XX	
PA	(DRMA/) DRMANAC R T.
PA	(LABA/) LABAT I.
PA	(STAC/) STACHE-CRAIN B.
PA	(DICK/) DICKSON M C.
PA	(JONE/) JONES L W.
XX	
PI	Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX	
DR	WPI, 2003-615964/58.
XX	
PT	New polynucleotide sequences obtained from various cDNA libraries, useful
PT	as hybridization probes, as oligomers for PCR, for chromosome and gene
PT	mapping, in the recombinant production of protein, or in generating
PT	antisense DNA or RNA.
XX	
FS	Claim 1; SEQ ID NO 31705; 44pp; English.
XX	
CC	The invention relates to an isolated polynucleotide comprising any one of
CC	38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC	determined by the technique of SBH (sequencing by hybridisation). Also
CC	included is a purified polypeptide comprising a sequence corresponding to
CC	a reading frame of the novel polynucleotide. The nucleic acid sequences
CC	are useful in diagnostics as expressed sequence tags (EST) for
CC	identifying expressed genes or for physical mapping of the human genome,
CC	in forensics, in assessing biodiversity, or in identifying mutations
CC	responsible for genetic disorders and other traits. The nucleotide
CC	sequences are also useful as hybridisation probes, as oligomers for PCR,
CC	for chromosome and gene mapping, in the recombinant production of
CC	protein, or in generating antisense DNA or RNA. The purified polypeptide
CC	is useful for generating antibodies specific for it. The present sequence
CC	is one of the 38043 isolated cDNA/EST sequences. Note: the sequence data
CC	for this patent did not form part of the printed specification, but was
CC	obtained in electronic format directly from USPTO at
CC	seqdata.uspto.gov/sequence.html?docID=20030073623
XX	
SQ	Sequence 464 BP; 96 A; 136 C; 131 G; 96 T; 0 U; 5 Other;
Query Match 25.5%; Score 265; DB 9; Length 464;	
Best Local Similarity 77.6%; Pred. No. 1.7e-67;	
Matches 333; Conservative 0; Mismatches 95; Indels 1; Gaps 1;	

QY 126 GCGGGGATGGACAGCTGCTCTTAGAGATTGCTAACTGAGAGAACCCACGAGAGACTCT 185
 Db |||||
 QY 186 CAATACCAAGTTGATGCTCTTATGGGGAGAGAGTACTCAGCAGACATCTATGCAGG 245
 Db |||||
 QY 197 CCTTACCTGTGCTTGCCTTTGACAGCCATAGGAGGACTGCCCTGACATGTACCTTCC 156
 Db |||||
 QY 246 TTCAGTGCAGCCCATCTTAAGGCATCTTGTGGAAGGCGCAGAAATGCCAGTGTCTGCCTA 305
 Db |||||
 QY 257 TTGACCGTGGACCGTGCCTGCTGCTCTTTCGTAATAAGCAGGAGCTGCGCTA 216
 Db |||||
 QY 306 TGAACCCACAGAGCTGGGAAGACGACACAATGCTGGGCGAGCCCAAGCAACTGGGGT 365
 Db |||||
 QY 217 CGGTGCTGAGGAGCTGGCAATAC-ATCTCGATGCTTGGCAGCCAGAGCAACTGGGGT 275
 Db |||||
 QY 366 GATCCCGCGGCTCTCATGAGCCTCTGAGCTCAGAGGAGAGGTTCCGAGGCGG 425
 Db |||||
 QY 276 GATCCCGCGGCTCTCATGAGCCTCTGAGCTCAGAGGAGAGGTTCCGAGGCGG 335
 Db |||||
 QY 426 GCCATGGGCGCTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGAGGTATTAGA 485
 Db |||||
 QY 336 GCATGGGCGCTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGAGGTATTAGA 395
 Db |||||
 QY 486 CTTCTGGACCTCTGCTTGGGAGACCTGGTAAATCCGAGAGACTGCCGGGGGAATATCCT 545
 Db |||||
 QY 396 CTTCTGGACCTCTGCTTGGGAGACCTTGGAAATCCGAGAGACTGCCGGGGGAATATCCT 455
 Db |||||
 QY 546 GATTCGGG 554
 Db |||||
 QY 456 GATTCGGG 464

RESULT 13

ACH77408
 ID ACH77408 standard; DNA; 531 BP.
 AC ACH77408;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Human genome derived single exon probe #10603.
 XX
 KW Human; probe; ss; gene expression; single exon probe; microarray;
 KW alternative splicing event; genomic alteration.
 XX
 OS Homo sapiens.
 XX
 PN US2003194704-A1.
 XX
 PD 16-OCT-2003.
 XX
 PF 03-APR-2002; 2002US-00029386.
 XX
 PR 03-APR-2002; 2002US-00029386.
 XX
 PA (PENN/) PENN S G.
 PA (RANK/) RANK D R.
 PA (HANZ/) HANZEL D K.
 XX
 PI Penn SG, Rank DR, Hanzel DK;
 XX
 DR WPI; 2004-119264/12.
 XX
 PT New human genome-derived single exon nucleic acid probes useful for human
 PT gene expression analysis, for identifying or characterizing alternative
 PT splicing events, for assessing genomic alterations or as tools for
 PT surveying tissues.
 XX
 PS Claim 15; SEQ ID NO 10603; 80pp; English.
 XX
 CC The invention relates to a nucleic acid probe for measuring human gene.

expression, comprising any of the 27,400 fully defined nucleotide
 sequences in the specification, or their complements or fragments, and
 encoding at least 8 amino acids of any of the 688 amino acid sequences
 fully defined in the specification. The probe is a single exon probe that
 hybridises under high stringency conditions to a nucleic acid molecule
 expressed in human cells or tissues. Also included are a spatially-
 addressable set of single exon nucleic acid probes for measuring human
 gene expression (comprising a plurality of single exon nucleic acid
 probes cited above, where each of the plurality of probes is separately
 and addressably isolatable or amplifiable from the plurality), a single
 exon microarray for measuring human gene expression, a method of
 measuring human gene expression, a vector comprising the single exon
 probe cited above, an ORF-encoded peptide comprising at least 8
 contiguous amino acids of any of the above-mentioned amino acid
 sequences (optionally with conservative amino acid substitutions), an
 isolated antibody that binds specifically to a peptide cited above,
 methods of selling and/or licensing single exon probes or microarrays to
 a customer desiring to measure gene expression, a method of providing
 human gene expression data by subscription, and a computer-readable
 storage medium which contains a database having a plurality of records
 (each record including data on the expression of a single exon probe
 cited above. The probe, methods and apparatus are useful in gene
 expression analysis. The probes may be used as tools for surveying
 tissues to detect the presence of expressed messages that contain their
 specific exon, or in constructing genome-derived single exon microarrays.
 In addition, the probes are used in identifying and characterising
 alternative splicing events, in detecting and characterising gross
 alterations in the genomic locus that includes their exon, in assessing
 smaller genomic alterations, in priming the synthesis of nucleic acids,
 or in expressing the ORF-encoded peptide. The present sequence is a human
 single exon probe of the invention. Note: The sequence data for this
 patent did not form part of the printed specification, but was obtained
 in electronic format directly from USPTO at
 cc seqdata.uspto.gov/sequence.html?DocID=20030194704
 XX
 SQ Sequence 531 BP; 117 A; 153 C; 149 G; 112 T; 0 U; 0 Other;
 Query Match 22.4%; Score 233.2; DB 12; Length 531;
 Best Local Similarity 98.7%; Pred. No. 4.2e-58;
 Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 684 CAAGGTGGACCGCGGAAACGTTTGGCCCCCATTTGCCAGCGAGAGGAAAATCTACCT 743
 Db |||||
 QY 34 CCAGGTGGACCGCGGAAACGTTTGGCCCCCATTTGCCAGCGAGAGGAAAATCTACCT 93
 Db |||||
 QY 744 GATTGACTTGGCTGGGTGAGAGGACAAACCGCGCACAGGCAACAGGGCTTCGGCTAAA 803
 Db |||||
 QY 94 GATTGACTTGGCTGGGTGAGAGGACAAACCGCGCACAGGCAACAGAGCCCTTCGGCTAAA 153
 Db |||||
 QY 804 AGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAGTGGTAGATCGGCTGAA 863
 Db |||||
 QY 154 AGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAGTGGTAGATCGGCTGAA 213
 Db |||||
 QY 864 TCAGGGCTCCCTCGGTGTTACCTTATCGGACAGCAAGCTCCTCGGCTATTGCGAGAC 921
 Db |||||
 QY 214 TCAGGGCTCCCTCGGTGTTACCTTATCGGACAGCAAGCTCCTCGGCTATTGCGAGTC 271
 Db |||||
 RESULT 14
 ACH91108
 ID ACH91108 standard; DNA; 232 BP.
 AC ACH91108;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Human genome derived single exon probe #24303.
 XX
 KW Human; probe; ss; gene expression; single exon probe; microarray;
 KW alternative splicing event; genomic alteration.
 XX
 OS Homo sapiens.
 XX


```

Db 406 CCACAGGCTGTGGGAAACCTACACATGCTGGGCACAGACCAGAGGCTTGCATCTAT 465
Qy 370 CCGCGGGCTCTCATGAGACCTCTGACAGCTCAAAAGGAGGAGGCTGCCAGGGCCGGCCA 429
Db 466 GTTCAGACCTCAACGACCTCTCCGTGCCATCGAGGAGACCAGCAATGA-----C 516
Qy 430 TGGGCCCTTCTGTGTCACCATGCTTTACCTAGAGATCTACAGGAGAGGTATTTAGACCTC 489
Db 517 ATGGAGTATGAGGTCTCCATGCTCTACCTGGAGATCTACAAATGAGATGATCCGGACCTG 576
Qy 490 CTGGACCTGCTTCGGGAGACCTGTAATCCGAGAGAGCTGCCGGGGAAATATCTGATT 549
Db 577 CTGAACCTCTCCCTGGCTACCTGGAGCTCGGGAGGACTCTAAGGGGGGTGATCCAGGTG 636
Qy 550 CCGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTGATTTTGAGCGGCACTTCTCGCCA 609
Db 637 GCGGCATCACGGAAGTCTCCACCATCAATGCCAGAGATCATGAGCTGTGATGAAG 696
Qy 610 GCCAGTCGAATCGGACTGTAGAGCCACC CGGCTCAACAGCGCTCTCTCCGCACTCAT 669
Db 697 GGGAAACCGCAGAGGACCAGGAGCCACCGCGGCCAACAGAGCTCTCTCCGCTCCAC 756
Qy 670 GCTGTCTCTCTGTCAGGTGAGCAGCGGAGACGTTTGGCCCCATTTGCCA-----G 723
Db 757 GGGTACTCAGGTGACCGTGGCCAGCGCAGCGGGTCAAGAAACATCTTGAGGAGGTG 816
Qy 724 CGAGAGGAAACTCTACCTGATTGCTTTGGTGGGTGAGAGCAACCGGCGCACAGGC 783
Db 817 CGGACGGCGGCTGTTCATGATCGACCTGGCTGGCTCAGAGCGGCTCGGAGACAG 876
Qy 784 AACAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTGCTCTGGGC 843
Db 877 AATCGTGGCAGCGTATGAAGAGGGGGCCACATCAACCGCTCACTGCTGGCACTGGGC 936
Qy 844 AAGTGTAGATGCGCTG-----AATCAGGGCTCCCTCGTGTACCTTATCGGACAGC 897
Db 937 AATGTCATCAACGCCCTGAGCGACAGAGGGTAGCAACAAGTACATCAACTATCGGACAGC 996
Qy 898 AAGCTCACCTGCTATTGAGAGCTCTGCGGTGGCTCAGCCACAGTATCCTATTGCC 957
Db 997 AAGCTCACCGGCTCTGAGAGCTCTCTGGAGGAAACAGCCGACAGTATGATCGCT 1056
Qy 958 AACATTGCCCCGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCC 1017
Db 1057 CACATCAGTCTCTGAGCAGTGCCTTCGAGGAGTCCGGAAACACCTGACCTACGCGGC 1116
Qy 1018 AGGTCAAGGAGGTGATCAATTG 1040
Db 1117 CGGGCCAGAACATTAAGACTAG 1139

```

Search completed: November 10, 2004, 06:57:56
 Job time : 541.26 secs

mis Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 13:18:18 ; Search time 555.617 Seconds
(without alignments)
10087.639 Million cell updates/sec

Title: US-10-797-893-5
Perfect score: 1041

Sequence: 1 atgggtcgctgtcgctaag.....ccaaggaggatgaattga 1041

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3611042 seqs, 2692057975 residues

Total number of hits satisfying chosen parameters: 7222084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10_NEW PUB.seq:
19: /cgn2_6/ptodata/1/pubpna/US11_NEW PUB.seq:
20: /cgn2_6/ptodata/1/pubpna/US60_NEW PUB.seq:
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1041	100.0	1041	US-10-797-893-5	Sequence 5, Appli
2	1038.4	99.8	1464	US-10-797-893-7	Sequence 7, Appli
3	1038	99.7	1115	US-10-797-893-1	Sequence 1, Appli
4	1035.4	99.5	1538	US-10-334-143-102	Sequence 3, Appli
5	1035.4	99.5	2099	US-10-334-143-102	Sequence 102, App
6	1034	99.3	2104	US-09-925-300-288	Sequence 288, App
7	355.4	34.1	24525	US-10-087-192-1180	Sequence 1180, Ap
8	265	25.5	464	US-09-918-995-31705	Sequence 31705, A
9	262.4	25.2	24291	US-10-087-192-1177	Sequence 1177, Ap
10	233.2	22.4	531	US-10-029-386-10603	Sequence 10603, A
11	230.4	22.1	232	US-10-029-386-24303	Sequence 24303, A
12	184.6	17.7	2095	US-10-108-260A-249	Sequence 249, App

13	155.8	15.0	3374	16	US-10-287-226-301	Sequence 301, App
14	136.6	13.1	1152	9	US-09-883-096-6	Sequence 6, Appli
15	136.6	13.1	4108	9	US-09-883-096-1	Sequence 1, Appli
16	135.2	13.0	2675	16	US-10-287-226-305	Sequence 305, App
17	129	12.4	2607	17	US-10-437-963-1460	Sequence 1460, Ap
18	127.6	12.3	482	16	US-10-152-319A-498	Sequence 498, App
19	123.2	11.8	2698	16	US-10-425-114-25129	Sequence 25129, A
20	121.6	11.7	3034	18	US-10-425-115-148256	Sequence 148256, A
21	121.4	11.7	1014	9	US-09-883-096-3	Sequence 3, Appli
22	114.6	11.0	3661	16	US-10-311-642-1	Sequence 1, Appli
23	113.6	10.9	987	16	US-10-311-642-3	Sequence 3, Appli
24	110.8	10.6	4292	16	US-10-287-226-313	Sequence 313, App
25	105.8	10.2	1998	17	US-10-437-963-56548	Sequence 56548, A
26	103.4	9.9	491	9	US-09-960-253-131	Sequence 131, App
27	101	9.7	2264	18	US-10-425-115-113546	Sequence 13546, A
28	100.2	9.6	3657	16	US-10-336-472-23	Sequence 23, Appli
29	100.2	9.6	3657	16	US-10-236-417-57	Sequence 57, Appli
30	94.2	9.0	4724	15	US-10-172-118-1087	Sequence 1087, Ap
31	94.2	9.0	4724	16	US-10-342-887-1087	Sequence 1087, Ap
32	93.8	9.0	3258	16	US-10-302-172-586	Sequence 586, App
33	93.4	9.0	2217	15	US-10-104-047-201	Sequence 201, App
34	90.4	8.7	4913	15	US-10-133-937-34	Sequence 34, Appli
35	90.4	8.7	4913	16	US-10-159-563-34	Sequence 27001, A
36	88.4	8.5	1295	15	US-10-104-047-762	Sequence 762, App
37	88.4	8.5	3159	15	US-10-425-114-22135	Sequence 22135, A
38	88	8.5	3390	16	US-10-739-930-2359	Sequence 2359, Ap
39	88	8.5	3407	18	US-10-425-115-46568	Sequence 46568, A
40	88	8.5	3433	18	US-10-172-118-1053	Sequence 1053, Ap
41	86.2	8.3	6927	15	US-10-342-887-1053	Sequence 1053, Ap
42	86.2	8.3	6927	16	US-10-424-599-111387	Sequence 111387, A
43	81.8	7.9	379	16	US-10-287-226-327	Sequence 327, App
44	80.8	7.8	4037	16	US-10-159-151-7	Sequence 7, Appli
45	79.2	7.6	1032	14	US-10-159-151-7	

ALIGNMENTS

RESULT 1
US-10-797-893-5
; Sequence 5, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Berand, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/797,893
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Human
US-10-797-893-5

Query Match	100.0%	Score 1041;	DB 17;	Length 1041;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1041;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGGTGCTGTGCGGTAAGCAAGATTGGAGTACTCGTCGTCACCTCCAGCTCGCGTA	60	
DB	1	ATGGGTGCTGTGCGGTAAGCAAGATTGGAGTACTCGTCGTCACCTCCAGCTCGCGTA	60	
QY	61	AGGGTGGCTGTGCGACTCGCGCCATTGTGGATGGAACAGCGGAGCAAGTATGATCCCCC	120	
DB	61	AGGGTGGCTGTGCGACTCGCGCCATTGTGGATGGAACAGCGGAGCAAGTATGATCCCCC	120	


```
QY 121 TGTGTGCGGGCATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCCAGCAG 180
Db |||||
QY 121 TGTGTGCGGGCATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCCAGCAG 180
Db |||||
QY 181 ACTCTCAATACACAGTTTGTATGCTCTCTATGGGAGAGAGTACTCAGCAGACATCTAT 240
Db |||||
QY 181 ACTCTCAATACACAGTTTGTATGCTCTCTATGGGAGAGAGTACTCAGCAGACATCTAT 240
Db |||||
QY 241 GCAGTTTCAGTGCAGCCCATCTAAGGCACTTGTCTGGAAGGCAAGATGCGAGTGTCTT 300
Db |||||
QY 241 GCAGTTTCAGTGCAGCCCATCTAAGGCACTTGTCTGGAAGGCAAGATGCGAGTGTCTT 300
Db |||||
QY 301 GCCTATGACCCACAGGAGCTGGGAGAGCAGCACAATGCTGGGAGCCAGAGCAACCT 360
Db |||||
QY 301 GCCTATGACCCACAGGAGCTGGGAGAGCAGCACAATGCTGGGAGCCAGAGCAACCT 360
Db |||||
QY 361 GGGGTGATGTCGGGGCTCTCATGACCTCTGACAGCTCACAAGGAGGAGGTCGCGAG 420
Db |||||
QY 361 GGGGTGATGTCGGGGCTCTCATGACCTCTGACAGCTCACAAGGAGGAGGTCGCGAG 420
Db |||||
QY 421 GGCAGGATGAGGCTGCTTCTGTCACCATGCTTACCTAGAGATCTACAGAGAGGTA 480
Db |||||
QY 481 TTAGACCTCTGACCCCTCTTCTGTCACCATGCTTACCTAGAGATCTACAGAGAGGTA 480
Db |||||
QY 481 TTAGACCTCTGACCCCTCTTCTGTCACCATGCTTACCTAGAGATCTACAGAGAGGTA 480
Db |||||
QY 481 TTAGACCTCTGACCCCTCTTCTGTCACCATGCTTACCTAGAGATCTACAGAGAGGTA 480
Db |||||
QY 541 ATCTGATGTCGGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTGAGGGGCA 600
Db |||||
QY 541 ATCTGATGTCGGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTGAGGGGCA 600
Db |||||
QY 601 TTCCTGCGAGCAGTCGAAATCGAATCTGAGAGGCAACCGGCTCAACAGGCTCTCTCC 660
Db |||||
QY 601 TTCCTGCGAGCAGTCGAAATCGAATCTGAGAGGCAACCGGCTCAACAGGCTCTCTCC 660
Db |||||
QY 661 GCGAGTCATGCTGCTCTGCTGCTCAAGGTGAGACCGGAGACGTTTGGCCCATTTGCG 720
Db |||||
QY 661 GCGAGTCATGCTGCTCTGCTGCTCAAGGTGAGACCGGAGACGTTTGGCCCATTTGCG 720
Db |||||
QY 721 CAGCAGAGGAGGAAACTCTTACCTGATGACTTGGCTGAGAGCAACCGGCGCACA 780
Db |||||
QY 721 CAGCAGAGGAGGAAACTCTTACCTGATGACTTGGCTGAGAGCAACCGGCGCACA 780
Db |||||
QY 781 GGCACACAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGCTGCTG 840
Db |||||
QY 841 GGCACACAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGCTGCTG 840
Db |||||
QY 841 GGCACACAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGCTGCTG 840
Db |||||
QY 901 CTCACCTGCTATGAGAGCTCTCTGGTGGCTCAGGCCCTCCTGCTGATCTTATCGGAGCAGCAAG 900
Db |||||
QY 901 CTCACCTGCTATGAGAGCTCTCTGGTGGCTCAGGCCCTCCTGCTGATCTTATCGGAGCAGCAAG 900
Db |||||
QY 901 CTCACCTGCTATGAGAGCTCTCTGGTGGCTCAGGCCCTCCTGCTGATCTTATCGGAGCAGCAAG 900
Db |||||
QY 961 ATTGCCCTGAGAGAGCTTCTACCTAGACAGAGTCTCGGACTCAACTTTCGCTGCGAGG 1020
Db |||||
QY 961 ATTGCCCTGAGAGAGCTTCTACCTAGACAGAGTCTCGGACTCAACTTTCGCTGCGAGG 1020
Db |||||
QY 1021 TCCAGGAGGTGATCAATTGA 1041
Db |||||
QY 1021 TCCAGGAGGTGATCAATTGA 1041
Db |||||
```

RESULT 2

US-10-797-893-7

; Sequence 7, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use

```
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10797,893  
; CURRENT FILING DATE: 2004-03-09  
; PRIOR APPLICATION NUMBER: US/09/724,224  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/597,292  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 1464  
; TYPE: DNA  
; ORGANISM: Human  
US-10-797-893-7  
  
Query Match 99.8%; Score 1038.4; DB 17; Length 1464;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 ATGGGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTA 60  
Db |||||  
QY 1 ATGGGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTA 60  
Db |||||  
QY 61 AGGGTGGCTGTGCGACTGCGGCCATTTGTGGATGGAAACAGCGGGAGCAAGTGAATCCCC 120  
Db |||||  
QY 61 AGGGTGGCTGTGCGACTGCGGCCATTTGTGGATGGAAACAGCGGGAGCAAGTGAATCCCC 120  
Db |||||  
QY 121 TGTGTGCGGGCATGAGCAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCCAGCAG 180  
Db |||||  
QY 121 TGTGTGCGGGCATGAGCAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCCAGCAG 180  
Db |||||  
QY 181 ACTCTCAATACACAGTTTGTATGCTCTCTATGGGAGAGAGTACTCAGCAGACATCTAT 240  
Db |||||  
QY 181 ACTCTCAATACACAGTTTGTATGCTCTCTATGGGAGAGAGTACTCAGCAGACATCTAT 240  
Db |||||  
QY 241 GCAGTTTCAGTGCAGCCCATCTAAGGCACTTGTCTGGAAGGCAAGATGCGAGTGTCTT 300  
Db |||||  
QY 241 GCAGTTTCAGTGCAGCCCATCTAAGGCACTTGTCTGGAAGGCAAGATGCGAGTGTCTT 300  
Db |||||  
QY 301 GCCTATGACCCACAGGAGCTGGGAGAGCAGCACAATGCTGGGAGCCAGAGCAACCT 360  
Db |||||  
QY 301 GCCTATGACCCACAGGAGCTGGGAGAGCAGCACAATGCTGGGAGCCAGAGCAACCT 360  
Db |||||  
QY 361 GGGGTGATGTCGGGGCTCTCATGACCTCTGACAGCTCACAAGGAGGAGGTCGCGAG 420  
Db |||||  
QY 361 GGGGTGATGTCGGGGCTCTCATGACCTCTGACAGCTCACAAGGAGGAGGTCGCGAG 420  
Db |||||  
QY 421 GGCAGGATGAGGCTGCTTCTGTCACCATGCTTACCTAGAGATCTACAGAGAGGTA 480  
Db |||||  
QY 421 GGCAGGATGAGGCTGCTTCTGTCACCATGCTTACCTAGAGATCTACAGAGAGGTA 480  
Db |||||  
QY 481 TTAGACCTCTGAGACCTGCTTTCGGGAGACTGGTAAATCCGAGAGAGCTCGCGGGGAT 540  
Db |||||  
QY 481 TTAGACCTCTGAGACCTGCTTTCGGGAGACTGGTAAATCCGAGAGAGCTCGCGGGGAT 540  
Db |||||  
QY 541 ATCTGATGTCGGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTGAGGGGCA 600  
Db |||||  
QY 541 ATCTGATGTCGGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTGAGGGGCA 600  
Db |||||  
QY 601 TTCCTGCGAGCAGTCGAAATCGAATCTGAGAGGCAACCGGCTCAACAGGCTCTCTCC 660  
Db |||||  
QY 601 TTCCTGCGAGCAGTCGAAATCGAATCTGAGAGGCAACCGGCTCAACAGGCTCTCTCC 660  
Db |||||  
QY 661 GCGAGTCATGCTGCTCTGCTGCTCAAGGTGAGACCGGAGACGTTTGGCCCATTTGCG 720  
Db |||||  
QY 661 GCGAGTCATGCTGCTCTGCTGCTCAAGGTGAGACCGGAGACGTTTGGCCCATTTGCG 720  
Db |||||  
QY 721 CAGCAGAGGAGGAAACTCTTACCTGATGACTTGGCTGAGAGCAACCGGCGCACA 780  
Db |||||  
QY 721 CAGCAGAGGAGGAAACTCTTACCTGATGACTTGGCTGAGAGCAACCGGCGCACA 780  
Db |||||  
QY 781 GGCACACAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGCTGCTG 840  
Db |||||  
QY 781 GGCACACAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGCTGCTG 840  
Db |||||
```

```
Db 781 GGCACAAAGGGCCCTTGGCTGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTGTCCTG 840
Qy 841 GGCACAAAGTGTAGATGCGCTGAATCAGGGCTCCCTGCTGTACTTATTCGGGACAGCAAG 900
Db 841 GGCACAAAGTGTAGATGCGCTGAATCAGGGCTCCCTGCTGTACTTATTCGGGACAGCAAG 900
Qy 901 CTCACTCGCCCTATTGCAAGACTCTCTGGGTGGCTCAGCCACAGATATCTTATTGCAAC 960
Db 901 CTCACTCGCCCTATTGCAAGACTCTCTGGGTGGCTCAGCCACAGATATCTTATTGCAAC 960
Qy 961 ATTGCCCTCAGAGAGCTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGG 1020
Db 961 ATTGCCCTCAGAGAGCTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGG 1020
Qy 1021 TCCAAGGAGGTGATCAATTG 1040
Db 1021 TCCAAGGAGGTGATCAATCG 1040

RESULT 3
US-10-797-893-1
; Sequence 1, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/797,893
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1115
; TYPE: DNA
; ORGANISM: Human
US-10-797-893-1

Query Match 99.7%; Score 1038; DB 17; Length 1115;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GGTGCTGTGCGCTAGCAAGATTGGAGCTACTCGTGTCCACTCCAGTCGCGTAAGG 63
Db 78 GGTGCTGTGCGCTAGCAAGATTGGAGCTACTCGTGTCCACTCCAGTCGCGTAAGG 137
Qy 64 GTGGCTGTGCACTGCGGCCATTGTGGATGGAAACAGCGGGAGCAAGTGATCCCTGT 123
Db 138 GTGGCTGTGCACTGCGGCCATTGTGGATGGAAACAGCGGGAGCAAGTGATCCCTGT 197
Qy 124 GTGGGGGCAATGACAGCTCTCTAGAGATTGTCTAACTGGAGGAACACACAGGAGACT 183
Db 198 GTGGGGGCAATGACAGCTCTCTAGAGATTGTCTAACTGGAGGAACACACAGGAGACT 257
Qy 184 CTCAATACCAAGTTTGAATGCTCTATGGGGAGAGAGTACTCAGAGGACATCTATGCA 243
Db 258 CTCAATACCAAGTTTGAATGCTCTATGGGGAGAGAGTACTCAGAGGACATCTATGCA 317
Qy 244 GGTTCAGTCAGCCCTCAAGGCACTTCTGAGAGGGGAGAGTCCAGTGCTGTTGCC 303
Db 318 GGTTCAGTCAGCCCTCAAGGCACTTCTGAGAGGGGAGAGTCCAGTGCTGTTGCC 377
Qy 304 TATGGACCCACAGGAGCTGGGAAGAGCGCACACATGCTGGGAGCCCGAGCAACTCTGGG 363
Db 378 TATGGACCCACAGGAGCTGGGAAGAGCGCACACATGCTGGGAGCCCGAGCAACTCTGGG 437
Qy 364 GTGATCCCGGGCTCTCATGGAACCTCTGACGCTCAAGGAGGAGGTGCGGAGGC 423
Db 438 GTGATCCCGGGCTCTCATGGAACCTCTGACGCTCAAGGAGGAGGTGCGGAGGC 497
```

```
Qy 424 CGSCCATGGCCCTTTCTGTCAACATGTCTTACCTAGAGATCTACAGGAGAAAGTATTA 483
Db 498 CGSCCATGGCCCTTTCTGTCAACATGTCTTACCTAGAGATCTACAGGAGAAAGTATTA 557
Qy 484 GACCTCTGGACCTCTGCTCGGAGAGACTGGTAATCCGAGAAGA CTGCCGGGGAAATATC 543
Db 558 GACCTCTGGACCTCTGCTCGGAGAGACTGGTAATCCGAGAAGA CTGCCGGGGAAATATC 617
Qy 544 CTGATTCGGGGTCTCTCCAGAGAGCCATCAGTAGCTTTGCTGATTTTGAGCGGCACTTC 603
Db 618 CTGATTCGGGGTCTCTCCAGAGAGCCATCAGTAGCTTTGCTGATTTTGAGCGGCACTTC 677
Qy 604 CTGCCAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACAGAGCCTCTCTCCCGC 663
Db 678 CTGCCAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACAGAGCCTCTCTCCCGC 737
Qy 664 AGTCATGCTGTGCTCTCTGCTCAAGGTGGACCGGGAGAGCTTTGGCCCCCATTTGCGCAG 723
Db 738 AGTCATGCTGTGCTCTCTGCTCAAGGTGGACCGGGAGAGCTTTGGCCCCCATTTGCGCAG 797
Qy 724 CGAGAGGGAAACTCTCTACCTGATTGACTTGGCTGGGTCTAGAGGACAAACCGCGCACAGGC 783
Db 798 CGAGAGGGAAACTCTCTACCTGATTGACTTGGCTGGGTCTAGAGGACAAACCGCGCACAGGC 857
Qy 784 AACAGGGCTCTCGGCTAAAAGAGAGTGGAGCCATCAACACCTCTCTGTTGCTCTGGGC 843
Db 858 AACAGGGCTCTCGGCTAAAAGAGAGTGGAGCCATCAACACCTCTCTGTTGCTCTGGGC 917
Qy 844 AAGTGTGATGCGCTGAATCAGGGCTCCCTCGTGTACCTTATCGGAGCAGCAAGCTC 903
Db 918 AAGTGTGATGCGCTGAATCAGGGCTCCCTCGTGTACCTTATCGGAGCAGCAAGCTC 977
Qy 904 ACTCGCTATTGCGAGCACTCTCTGGGTGGCTCAGGCCACAGTATCTTATTGCCAACATT 963
Db 978 ACTCGCTATTGCGAGCACTCTCTGGGTGGCTCAGGCCACAGTATCTTATTGCCAACATT 1037
Qy 964 GCGCTGTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTGTGTCAGAGTCC 1023
Db 1038 GCGCTGTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTGTGTCAGAGTCC 1097
Qy 1024 AAGAGGTGATCAATTGA 1041
Db 1098 AAGAGGTGATCAATTGA 1115
```

RESULT 4

```
US-10-797-893-3
; Sequence 3, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/797,893
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: Human
US-10-797-893-3
```

Query Match 99.5%; Score 1035.4; DB 17; Length 1538;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG 63
 DB |||||
 QY 78 GGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG 137
 DB |||||
 QY 64 GTGGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG 123
 DB |||||
 QY 138 GTGGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG 197
 DB |||||
 QY 124 GTGGGGGATGAGACAGCTGCTCTCTAGAGATTGCTAATCGAGAGAACCCAGAGAGACT 183
 DB |||||
 QY 198 GTGGGGGATGAGACAGCTGCTCTCTAGAGATTGCTAATCGAGAGAACCCAGAGAGACT 257
 DB |||||
 QY 184 CTCGAATACAGTTGATGCTCTCTAGAGATTGCTAATCGAGAGAACCCAGAGAGACT 243
 DB |||||
 QY 258 CTCGAATACAGTTGATGCTCTCTAGAGATTGCTAATCGAGAGAACCCAGAGAGACT 317
 DB |||||
 QY 244 GGTTCAGTGCAGCCATCTTAAGGCACTTGTCTGGAAGGCGAGATGCTGCTGCTGCC 303
 DB |||||
 QY 318 GGTTCAGTGCAGCCATCTTAAGGCACTTGTCTGGAAGGCGAGATGCTGCTGCTGCC 377
 DB |||||
 QY 304 TATGAGCCACAGGAGCTGGAGAGCGCACAAATGCTGGGAGGCCAGAGCAACTGGG 363
 DB |||||
 QY 378 TATGAGCCACAGGAGCTGGAGAGCGCACAAATGCTGGGAGGCCAGAGCAACTGGG 437
 DB |||||
 QY 364 GTGATCCCGGGCTCTCATGACCTCTGACAGCTCACAAAGGAGGAGGTCGCGAGGGC 423
 DB |||||
 QY 438 GTGATCCCGGGCTCTCATGACCTCTGACAGCTCACAAAGGAGGAGGTCGCGAGGGC 497
 DB |||||
 QY 424 CGGCCATGGGCCCTTCTGTCAACCATGCTTACCTAGAGATCTACAGAGAGAGTATTA 483
 DB |||||
 QY 498 CGGCCATGGGCCCTTCTGTCAACCATGCTTACCTAGAGATCTACAGAGAGAGTATTA 557
 DB |||||
 QY 484 GACCTCTGGAACCTGCTTGGGAGAGCTGTAATCGAGAGAGACTGCGGGGGAATATC 543
 DB |||||
 QY 558 GACCTCTGGAACCTGCTTGGGAGAGCTGTAATCGAGAGAGACTGCGGGGGAATATC 617
 DB |||||
 QY 544 CTGATTCGGGCTCTCCAGAGCCCATCAGTACGTTTGTGATTTTGGAGGGGACTTC 603
 DB |||||
 QY 618 CTGATTCGGGCTCTCCAGAGCCCATCAGTACGTTTGTGATTTTGGAGGGGACTTC 677
 DB |||||
 QY 604 CTGCCAGCAGTCGAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCCCGC 663
 DB |||||
 QY 678 CTGCCAGCAGTCGAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCCCGC 737
 DB |||||
 QY 664 AGTCATGCTGCTCTGCTCAAGTGGAACAGCGGGAACGTTTGGGCCCCCATTTGCGCAG 723
 DB |||||
 QY 738 AGTCATGCTGCTCTGCTCAAGTGGAACAGCGGGAACGTTTGGGCCCCCATTTGCGCAG 797
 DB |||||
 QY 724 CGAGAGGGAACCTCTACCTGATTGACTTGGCTGAGAGCAACCCGGGCGCAGGC 783
 DB |||||
 QY 798 CGAGAGGGAACCTCTACCTGATTGACTTGGCTGAGAGCAACCCGGGCGCAGGC 857
 DB |||||
 QY 784 AACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTCTCTGGGC 843
 DB |||||
 QY 858 AACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTCTCTGGGC 917
 DB |||||
 QY 844 AAGTGTGATGAGTGGTGAATCAGGGCTCTCTGTTGATCTTATCGGGAGCAGAGCTC 903
 DB |||||
 QY 918 AAGTGTGATGAGTGGTGAATCAGGGCTCTCTGTTGATCTTATCGGGAGCAGAGCTC 977
 DB |||||
 QY 904 ACTCGCTATTGAGAGCTCTCTGGTGGCTCAGCCCAAGTATCTTTATTCGCAACTT 963
 DB |||||
 QY 978 ACTCGCTATTGAGAGCTCTCTGGTGGCTCAGCCCAAGTATCTTTATTCGCAACTT 1037
 DB |||||
 QY 964 GCCCTCTGAGAGAGCTCTTACCTAGACAGCTCTCGGACTCAACTTCTCTCGAGGTCC 1023
 DB |||||
 QY 1038 GCCCTCTGAGAGAGCTCTTACCTAGACAGCTCTCGGACTCAACTTCTCTCGAGGTCC 1097
 DB |||||
 QY 1024 AAGAGGCTGATCAATTG 1040
 DB |||||
 QY 1098 AAGAGGCTGATCAATCG 1114
 DB |||||

RESULT 5
 US-10-334-143-102
 ; Sequence 102, Application US/10334143
 ; Publication No. US20040009549A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH
 ; APPLICANT: SUDARSANAM, SUCHA
 ; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
 ; FILE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD
 ; FILE REFERENCE: 038602/1543
 ; CURRENT APPLICATION NUMBER: US/10/334,143
 ; CURRENT FILING DATE: 2002-12-31
 ; PRIOR APPLICATION NUMBER: 60/343,169
 ; PRIOR FILING DATE: 2001-12-31
 ; NUMBER OF SEQ ID NOS: 207
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 102
 ; LENGTH: 2099
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-334-143-102

 Query Match 99.5%; Score 1035.4; DB 16; Length 2099;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 QY 4 GGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG 63
 DB |||||
 QY 101 GGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG 160
 DB |||||
 QY 64 GTGGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG 123
 DB |||||
 QY 161 GTGGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG 220
 DB |||||
 QY 124 GTGGGGGATGAGACAGCTGCTCTCTAGAGATTGCTAATCGAGAGAACCCAGAGAGACT 183
 DB |||||
 QY 221 GTGGGGGATGAGACAGCTGCTCTCTAGAGATTGCTAATCGAGAGAACCCAGAGAGACT 280
 DB |||||
 QY 184 CTCGAATACAGTTGATGCTCTCTAGAGATTGCTAATCGAGAGAACCCAGAGAGACT 243
 DB |||||
 QY 281 CTCGAATACAGTTGATGCTCTCTAGAGATTGCTAATCGAGAGAACCCAGAGAGACT 340
 DB |||||
 QY 244 GTTCAGTGCAGCCATCTTAAGGCACTTGTCTGGAAGGCGAGATGCTGCTGCTGCC 303
 DB |||||
 QY 341 GTTCAGTGCAGCCATCTTAAGGCACTTGTCTGGAAGGCGAGATGCTGCTGCTGCC 400
 DB |||||
 QY 304 TATGAGCCACAGGAGCTGGGAGAGCGCACAAATGCTGGGAGGCCAGAGCAACTGGG 363
 DB |||||
 QY 401 TATGAGCCACAGGAGCTGGGAGAGCGCACAAATGCTGGGAGGCCAGAGCAACTGGG 460
 DB |||||
 QY 364 GTGATCCCGGGCTCTCATGACCTCTGACAGCTCACAAAGGAGGAGGTCGCGAGGGC 423
 DB |||||
 QY 461 GTGATCCCGGGCTCTCATGACCTCTGACAGCTCACAAAGGAGGAGGTCGCGAGGGC 520
 DB |||||
 QY 424 CGGCCATGGGCCCTTCTGTCAACCATGCTTACCTAGAGATCTACAGAGAGAGTATTA 483
 DB |||||
 QY 521 CGGCCATGGGCCCTTCTGTCAACCATGCTTACCTAGAGATCTACAGAGAGAGTATTA 580
 DB |||||
 QY 484 GACCTCTGGAACCTGCTTGGGAGAGCTGTAATCGAGAGAGACTGCGGGGGAATATC 543
 DB |||||
 QY 581 GACCTCTGGAACCTGCTTGGGAGAGCTGTAATCGAGAGAGACTGCGGGGGAATATC 640
 DB |||||
 QY 544 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTATCTTGTGATTTTGGAGGGGACTTC 603
 DB |||||
 QY 641 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTATCTTGTGATTTTGGAGGGGACTTC 700
 DB |||||
 QY 604 CTGCCAGCAGTCGAATCGGAAATCGGAGCTTAGAGCCACCCGGCTCAACAGCGCTCTCCCGC 663
 DB |||||
 QY 701 CTGCCAGCAGTCGAATCGGAAATCGGAGCTTAGAGCCACCCGGCTCAACAGCGCTCTCCCGC 760
 DB |||||
 QY 664 AGTCATGCTGCTCTCTGCTCAAGTGGAACAGCGGGAACGTTTGGGCCCCCATTTGCCAG 723
 DB |||||
 QY 761 AGTCATGCTGCTCTCTGCTCAAGTGGAACAGCGGGAACGTTTGGGCCCCCATTTGCCAG 820
 DB |||||

QY 724 CGAGAGGGAATACTCTACCTGATTGACTTGGCTGGGTGAGAGCAACCGGCGCAGAGC 783
 DB 821 CGAGAGGGAATACTCTACCTGATTGACTTGGCTGGGTGAGAGCAACCGGCGCAGAGC 880
 QY 784 AACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTGCTTGGGC 843
 DB 881 AACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTGCTTGGGC 940
 QY 844 AAGTGTAGATCGCTGAATCAGGGCTCCCTCGTGTACCTTATCGGACAGCAAGCTC 903
 DB 941 AAGTGTAGATCGCTGAATCAGGGCTCCCTCGTGTACCTTATCGGACAGCAAGCTC 1000
 QY 904 ACTGCGCTATTGAGAGACTCTCTGGGTGGCTCAGCCACAGTAGTATCTTATTGCAACATT 963
 DB 1001 ACTGCGCTATTGAGAGACTCTCTGGGTGGCTCAGCCACAGTAGTATCTTATTGCAACATT 1060
 QY 964 GCGGCTGAGAGAGCTTCTACCTAGACACAGTCTCGGACTCAACTTGGTGGCAGGTCC 1023
 DB 1061 GCGGCTGAGAGAGCTTCTACCTAGACACAGTCTCGGACTCAACTTGGTGGCAGGTCC 1120
 QY 1024 AAGAGGTGATCAATTG 1040
 DB 1121 AAGAGGTGATCAATCG 1137

RESULT 6

US-09-925-300-288
 ; Sequence 288, Application US/09925300
 ; Patent No. US20020151681A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Craig Rosen,
 ; APPLICANT: Steve Ruben
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA101
 ; CURRENT APPLICATION NUMBER: US/09/925,300
 ; PRIOR FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05988
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 1890
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 288
 ; LENGTH: 2104
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (44)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: misc feature
 ; LOCATION: (497)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: misc feature
 ; LOCATION: (1323)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; US-09-925-300-288

Query Match 99.3%; Score 1034; DB 9; Length 2104;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1034; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 GGTGCTGTCGGCTAGCAGATTGGAGTACTCGTCCACCTCCAGCTCGCGTAAGG 63
 DB 67 GGTGCTGTCGGCTAAGCAAGATTGGAGTACTCGTCCACCTCCAGCTCGCGTAAGG 126
 QY 64 GTGGCTGTCGCACTCGGCGCAATTGTGGATGGAACAGCGGGAGCAAGTATCCCCCTGT 123
 DB 127 GTGGCTGTCGCACTCGGCGCAATTGTGGATGGAACAGCGGGAGCAAGTATCCCCCTGT 186
 QY 124 GTGCGGGGATCGACAGCTGCTCTTAGAGATTGCTTAAGTGGAGAACACACAGGAGACT 183

DB 187 GTGCGGGGATCGACAGCTGCTCTTAGAGATTGCTAACTGGAGGAACACACAGGAGACT 246
 QY 184 CTCAATACACAGTTTGTATGCTTCTATGSGGAGAGAGTACTCAGCAGGAGACATCTATGCA 243
 DB 247 CTCAATACACAGTTTGTATGCTTCTATGSGGAGAGAGTACTCAGCAGGAGACATCTATGCA 306
 QY 244 GGTTCAGTGCAGCCCATCTTAAGGCATTTGCTGGAAGGGCAGAAATCCAGTGTGCTTGGC 303
 DB 307 GGTTCAGTGCAGCCCATCTTAAGGCATTTGCTGGAAGGGCAGAAATCCAGTGTGCTTGGC 366
 QY 304 TATGAGCCACAGGAGCTGGGAAGAGCGACACAACTGCTGGGAGCCCGCAGAGCAACCTGGG 363
 DB 367 TATGAGCCACAGGAGCTGGGAAGAGCGACACAACTGCTGGGAGCCCGCAGAGCAACCTGGG 426
 QY 364 GTGATCCCGCGGCTCTCATGACCTCTGAGCTCACAAGGGAGGAGGTGCGGAGGGC 423
 DB 427 GTGATCCCGCGGCTCTCATGACCTCTGAGCTCACAAGGGAGGAGGTGCGGAGGGC 486
 QY 424 CGGCCATGGGCGCTTCTGTCAACCATGTCTTACCTAGAGATCTTACAGGAGAGGTATTATTA 483
 DB 487 CGGCCATGGGCGCTTCTGTCAACCATGTCTTACCTAGAGATCTTACAGGAGAGGTATTATTA 546
 QY 484 GACCTCTGGAGCCCTGCTTCGGGAGACCTTGGTAAATCCGAGAGAGACTGCGGCGGGAATATC 543
 DB 547 GACCTCTGGAGCCCTGCTTCGGGAGACCTTGGTAAATCCGAGAGAGACTGCGGCGGGAATATC 606
 QY 544 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTTCGAGGGGCACTTC 603
 DB 607 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTTCGAGGGGCACTTC 666
 QY 604 TTGCGCAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACACAGGCGCTCTCCCGC 663
 DB 667 TTGCGCAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACACAGGCGCTCTCCCGC 726
 QY 664 AGTCATGCTGCTCTCTGCTCAAGGTGACACAGGGGAAACGTTTGGCCCCCATTTTGGCCAG 723
 DB 727 AGTCATGCTGCTCTCTGCTCAAGGTGACACAGGGGAAACGTTTGGCCCCCATTTTGGCCAG 786
 QY 724 CGAGAGGGGAAACACTCTACCTGATTGACTTTGGCTGAGAGAGCAACCGGCGCAGAGC 783
 DB 787 CGAGAGGGGAAACACTCTACCTGATTGACTTTGGCTGAGAGAGCAACCGGCGCAGAGC 846
 QY 784 AACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTGTTTGTCTTGGGC 843
 DB 847 AACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTGTTTGTCTTGGGC 906
 QY 844 AAGTGTAGATCGCTGAATCAGGGCTCTCTCGTGTACCTTATCGGAGCAAGAGCTC 903
 DB 907 AAGTGTAGATCGCTGAATCAGGGCTCTCTCGTGTACCTTATCGGAGCAAGAGCTC 966
 QY 904 ACTCGCTTATTCAGGAGCTCTCTGGTGGCTCAGCCCGCAGTATCTTATTCGCAACATT 963
 DB 967 ACTCGCTTATTCAGGAGCTCTCTGGTGGCTCAGCCCGCAGTATCTTATTCGCAACATT 1026
 QY 964 GCGGCTGAGAGAGCTTCTTACCTAGACAGCTCTCGGACTCAACTTTCTGCGAGGTCC 1023
 DB 1027 GCGGCTGAGAGAGCTTCTTACCTAGACAGCTCTCGGACTCAACTTTCTGCGAGGTCC 1086
 QY 1024 AAGAGGTGATCAATTG 1040
 DB 1087 AAGAGGTGATCAATCG 1103

RESULT 7

US-10-087-192-1180
 ; Sequence 1180, Application US/10087192
 ; Publication No. US20020182586A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Engelhard, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 ; TITLE OF INVENTION: CANCER
 ; FILE REFERENCE: 529452000122

```

; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1180
; LENGTH: 24525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1180

Query Match      34.1%; Score 355.4; DB 13; Length 24525;
Best Local Similarity 84.8%; Pred. No. 2.3e-103;
Matches 445; Conservative 0; Mismatches 1; Indels 79; Gaps 1;

QY 476 AGGTATTAGACCTCTCGGACCTGCTCGGAGACCTGTTAATCCGAGAAGACTGCCGGG 535
DB 2313 AGGTATTAGACCTCTCGGACCTGCTCGGAGACCTGTTAATCCGAGAAGACTGCCGGG 2372
QY 536 GGAATATCTGATTCGGGTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGAGC 595
DB 2373 GGAATATCTGATTCGGGTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGAGC 2432
QY 596 GGCACCTCTGCGACCGAGTGAATCGGACTGTAGGAGCCACCGGCTCAACCGGCT 655
DB 2433 GGCACCTCTGCGACCGAGTGAATCGGACTGTAGGAGCCACCGGCTCAACCGGCT 2492
QY 656 CTTCCCGCAGTCACTGCTGCTCTGTTCA-----AGGTGACACG 685
DB 2493 CTTCCCGCAGTCACTGCTGCTCTGTTCAAGGTGAGCCGCGAGAGGAGGACCT 2552
QY 686 -----AGGTGACACG 696
DB 2553 GGAAGCCAGGAGCTGAGCTAAGCAGGAGACCTTTGTTCTTACCCCGAGTGGACCG 2612
QY 697 CGGGAACCTTTGGCCCAATTTGCGCAGCGAGAGGGAACCTTACCTGATTTGCTTGGCT 756
DB 2613 CGGGAACCTTTGGCCCAATTTGCGCAGCGAGAGGGAACCTTACCTGATTTGCTTGGCT 2672
QY 757 GGGTCAGAGGACAAACCGCGCACAGGCAACAAAGGCGCTTCGGCTAAAGAGAGTGGAGCC 816
DB 2673 GGGTCAGAGGACAAACCGCGCACAGGCAACAAAGGCGCTTCGGCTAAAGAGAGTGGAGCC 2732
QY 817 ATCAACACCTCCCTGTTTGTCTCGGCAAGTGTGTAGATGCGCTGAATTCAGGCGCTCCCT 876
DB 2733 ATCAACACCTCCCTGTTTGTCTCGGCAAGTGTGTAGATGCGCTGAATTCAGGCGCTCCCT 2792
QY 877 CGGTACCTTATCGGAGCAGCAAGCTCACTCGCTATTGACGAC 921
DB 2793 CGGTACCTTATCGGAGCAGCAAGCTCACTCGCTATTGACGATC 2837

```

```

RESULT 8
US-09-918-995-31705
; Sequence 31705, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31705
; LENGTH: 464
; TYPE: DNA

```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(464)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-31705

Query Match      25.5%; Score 265; DB 10; Length 464;
Best Local Similarity 77.6%; Pred. No. 1e-74;
Matches 333; Conservative 0; Mismatches 95; Indels 1; Gaps 1;

QY 126 GCGGGCATGACAGCTGCTCTTAGAGATTGCTAACTTGGAGGACCAACACAGGAGACTCT 185
DB 37 GCGTGGAAATCAATATCGCGTCAATAGTGTAAACGCCCCGAAACCAACCCGATAATCT 96
QY 186 CAAATACCAAGTTTGATGCTTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCAGG 245
DB 97 CCGTTACTGTGCTTGCCTTTTGACAGCCATAGGAGACTGCGCTGACATGTACCTTCC 156
QY 246 TTGAGTGCAGCCCATCTAAGGCACCTTGTGGAAGGGCAGAAATGCCAGTGTGCTTGCCCTA 305
DB 157 TTGACCGTGGACCGTGCACCTGCTCTTTTCGTAATAAAGCGAGGAGCTGGCCCTA 216
QY 306 TGGACCCACAGAGAGCTGGGAAGACGCACAAATCTGGGAGCCGACAGCAACCTGGGGT 365
DB 217 CGGTGCTCAGGAGCTGGCATTAC-ATCTGATGCTTGGCAGCCGACAGCAACCTGGGGT 275
QY 366 GATCCCGGGCTCTCATGGACCTCTGACCTCACAAGGAGGAGGAGGTGCCGAGGGCCG 425
DB 276 GATCCCGGGCTCTCATGGACCTCTGACCTCACAAGGAGGAGGAGGTGCCGAGGGCCG 335
QY 426 GCCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACGAGAGAGGTATTAGA 485
DB 336 GCCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACGAGAGAGGTATTAGA 395
QY 486 CCTCTGAGCCCTGCTTGGGAGAGCTGTAATCCGAGAAGACTGCCGGGGGAATATCCT 545
DB 396 CCTCTGAGCCCTGCTTGGGAGAGCTTGGNAATCCGAGAGAGACTGCCGGGGGAATATCCT 455
QY 546 GATTCGGG 554
DB 456 GATTCGGG 464

```

```

RESULT 9
US-10-087-192-1177
; Sequence 1177, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1177
; LENGTH: 24291
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(24291)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1177

```

```

Query Match      25.2%; Score 262.4; DB 13; Length 24291;

```



```

; TITLE OF INVENTION: No. US2004005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 249
; LENGTH: 2095
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-108-260A-249

Query Match          17.7%  Score 184.6;  DB 16;  Length 2095;
Best Local Similarity 53.7%;  Pred. No. 1.3e-48;
Matches 463;  Conservative 0;  Mismatches 379;  Indels 21;  Gaps 3;

QY 190 TACAGTTTATGCTTCTATGGGAGGAGTACTCAGCAGCAGCATCTATGTCAGGTTC 249
Db 190 TACAGTTTATGCTTCTATGGGAGGAGTACTCAGCAGCAGCATCTATGTCAGGTTC 249
QY 250 GTCCAGCCATCTAAGGCATCTTCTGGAAGGCGAATGCGCATGTGCTTGCCTATGGA 309
Db 250 GTCCAGCCATCTAAGGCATCTTCTGGAAGGCGAATGCGCATGTGCTTGCCTATGGA 309
QY 346 ACCAAGAGGCTCATCAGAGGCGTCTCAGGCTACAATGCCACTGTCTTGCCTATGCG 405
Db 346 ACCAAGAGGCTCATCAGAGGCGTCTCAGGCTACAATGCCACTGTCTTGCCTATGCG 405
QY 310 CCCACAGGAGCTGGGAAGCGCACAAATGCTGGGAGCCAGAGCAACTGGGGTGATC 369
Db 310 CCCACAGGAGCTGGGAAGCGCACAAATGCTGGGAGCCAGAGCAACTGGGGTGATC 369
QY 406 CCCACAGGCTGGGAAGCGCACAAATGCTGGGAGCCAGAGCAACTGGGGTGATC 465
Db 406 CCCACAGGCTGGGAAGCGCACAAATGCTGGGAGCCAGAGCAACTGGGGTGATC 465
QY 370 CCGCGGGCTCTCATGACCTCTTGCAGCTCACAAAGGAGGAGGGTCCGAGGGCCGGCCA 429
Db 370 CCGCGGGCTCTCATGACCTCTTGCAGCTCACAAAGGAGGAGGGTCCGAGGGCCGGCCA 429
QY 466 GTTCAGACCTCAACGACCTTTCCTGCTCATCGAGGAGCAGCAATGA-----C 516
Db 466 GTTCAGACCTCAACGACCTTTCCTGCTCATCGAGGAGCAGCAATGA-----C 516
QY 430 TGGGCGCTTCTGTACCATGCTTACCTAGAGATCTACAGGAGAGTATTAGACCTC 489
Db 430 TGGGCGCTTCTGTACCATGCTTACCTAGAGATCTACAGGAGAGTATTAGACCTC 489
QY 517 ATGGAGTATGAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 576
Db 517 ATGGAGTATGAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 576
QY 490 CTGGACCTCTCTCGGAGACCTGTGTAATCCGAGAGAGTCTCGGGGGAATATCTTGAT 549
Db 490 CTGGACCTCTCTCGGAGACCTGTGTAATCCGAGAGAGTCTCGGGGGAATATCTTGAT 549
QY 577 CTGAACCTCTCTCGGAGACCTGTGTAATCCGAGAGAGTCTCGGGGGAATATCTTGAT 636
Db 577 CTGAACCTCTCTCGGAGACCTGTGTAATCCGAGAGAGTCTCGGGGGAATATCTTGAT 636
QY 550 CCGGGTCTCTCCAGAGCCCATCAGTACGATGTTTGTGATTTTGTAGCGGCATCTTCCGCA 609
Db 550 CCGGGTCTCTCCAGAGCCCATCAGTACGATGTTTGTGATTTTGTAGCGGCATCTTCCGCA 609
QY 637 GCGGCGATCACCAGTCTCTCACCATCAATGCCAGGAGATCATGCGCTGCTGATGAAG 696
Db 637 GCGGCGATCACCAGTCTCTCACCATCAATGCCAGGAGATCATGCGCTGCTGATGAAG 696
QY 610 GCCAGTGCAGTCTGAGAGGCGCCGCTCAACGAGGCTCTCCCGGAGTAT 669
Db 610 GCCAGTGCAGTCTGAGAGGCGCCGCTCAACGAGGCTCTCCCGGAGTAT 669
QY 697 GGAACCGGAGAGGAGCCAGGAGGCGCCAGGAGGCGCCAGGAGGCGCCAGGAGGCGCC 756
Db 697 GGAACCGGAGAGGAGCCAGGAGGCGCCAGGAGGCGCCAGGAGGCGCCAGGAGGCGCC 756
QY 670 GCTGTGCTCTGTCTCAAGGTGAGCAGCGGAGAGTCTTGGCCCATTTCCGCA-----G 723
Db 670 GCTGTGCTCTGTCTCAAGGTGAGCAGCGGAGAGTCTTGGCCCATTTCCGCA-----G 723
QY 757 GCGGTACTGAGGTGAGTCTTGGGAGGAGTCTTGGGAGGAGTCTTGGGAGGAGTCTTGG 816
Db 757 GCGGTACTGAGGTGAGTCTTGGGAGGAGTCTTGGGAGGAGTCTTGGGAGGAGTCTTGG 816
QY 724 CGAGAGGGAATACTTACTGATGATCTTGGTGGGTGAGAGGAGCAACCGGCGCAGAGC 783
Db 724 CGAGAGGGAATACTTACTGATGATCTTGGTGGGTGAGAGGAGCAACCGGCGCAGAGC 783
QY 817 CGGAGGCGGCGCTGTTATGATCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 876
Db 817 CGGAGGCGGCGCTGTTATGATCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 876
QY 784 AACAGGCGCTTGGCTAAAGAGTGAAGGAGTCAACACCTCTCTTGTCTTGTCTTGGC 843
Db 784 AACAGGCGCTTGGCTAAAGAGTGAAGGAGTCAACACCTCTCTTGTCTTGTCTTGGC 843
QY 877 AATCGTGGGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 936
Db 877 AATCGTGGGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 936
QY 844 AAGTGGTATGATGCTG-----AATCAGGGGCTCTCTGCTGATCTTATCGGAGCAGC 897
Db 844 AAGTGGTATGATGCTG-----AATCAGGGGCTCTCTGCTGATCTTATCGGAGCAGC 897
QY 937 AACTGATCAACGCTCTGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 996
Db 937 AACTGATCAACGCTCTGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 996
QY 898 AAGTCACTCGCTATTCAGGAGTCTTGGTGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 957
Db 898 AAGTCACTCGCTATTCAGGAGTCTTGGTGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 957
QY 997 AAGCTCAGCGGCTCTTGAAGGAGTCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1056
Db 997 AAGCTCAGCGGCTCTTGAAGGAGTCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1056
QY 958 AACATTGCCCTGAGAGGCTTCTTACCTAGACAGTCTCGGAGTCTCGGAGTCTCGGAGTCT 1017
Db 958 AACATTGCCCTGAGAGGCTTCTTACCTAGACAGTCTCGGAGTCTCGGAGTCTCGGAGTCT 1017
QY 1057 CACATGAGTCTCGGAGGAGTCTTCTGAGGAGTCTTCTGAGGAGTCTTCTGAGGAGTCT 1116
Db 1057 CACATGAGTCTCGGAGGAGTCTTCTGAGGAGTCTTCTGAGGAGTCTTCTGAGGAGTCT 1116
QY 1018 AGGTCCAGGAGTATCAATTG 1040

```

```

Db 1117 CCGGCCAGACATTAAGACTAG 1139

RESULT 13
US-10-287-226-301
; Sequence 301, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
; APPLICANT: Agee, Michele L.,
; APPLICANT: Alsobrook, John P.,
; APPLICANT: Berchs, Constance,
; APPLICANT: Boldog, Ference,
; APPLICANT: Burgess, Catherine E.,
; APPLICANT: Chant, John S.,
; APPLICANT: Chaudhuri, Amitabha,
; APPLICANT: DiPippo, Vincent A.,
; APPLICANT: Edinger, Shlomit R.,
; APPLICANT: Eisen, Andrew,
; APPLICANT: Ellerman, Karen,
; APPLICANT: Gangolli, Esha A.,
; APPLICANT: Gorman, Linda,
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Khrantsov, Nikolai,
; APPLICANT: Li, Li,
; APPLICANT: Malyankar, Uriel M.,
; APPLICANT: MacDougall, John R.,
; APPLICANT: Mezes, Peter S.,
; APPLICANT: Miller, Charles E.,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Ooi, Chean Eng,
; APPLICANT: Ort, Tatiana,
; APPLICANT: Padigaru, Muralidhara,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rastelli, Luca,
; APPLICANT: Rieger, Daniel K.,
; APPLICANT: Rothenberg, Mark E.,
; APPLICANT: Shenoy, Suresh G.,
; APPLICANT: Spaderna, Steven K.,
; APPLICANT: Spytek, Kimberley A.,
; APPLICANT: Taupier, Jr., Raymond J.,
; APPLICANT: Vernet, Corine A.M.,
; APPLICANT: Zethusen, Bryan D.,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-480C
; CURRENT APPLICATION NUMBER: US/10/287,226
; PRIOR FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: 60/334,421
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,392
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/360,148
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 60/364,000
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/404,821
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/334,526
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,409
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/364,227
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/334,027
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/331,641
; PRIOR FILING DATE: 2001-11-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 673
; SOFTWARE: CuraSeqList version 0.1

```


; SEQ ID NO 301
 ; LENGTH: 3374
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (7)..(3351)
 US-10-287-226-301

Query Match 15.0%; Score 155.8; DB 16; Length 3374;
 Best Local Similarity 54.0%; Pred. No. 2.9e-39;
 Matches 394; Conservative 0; Mismatches 317; Indels 18; Gaps 3;

QY	321	TGGGAAGAGC	321	TGGGAAGAGC
DB	612	TGGGAAGAGC	612	TGGGAAGAGC
QY	381	CATGGACCTC	381	CATGGACCTC
DB	672	CAACGACCTC	672	CAACGACCTC
QY	441	TGTCACCATG	441	TGTCACCATG
DB	723	GGTCTCATG	723	GGTCTCATG
QY	501	TTCCGGAGAC	501	TTCCGGAGAC
DB	783	CCTGGGCTAC	783	CCTGGGCTAC
QY	561	CGAAGGCCAT	561	CGAAGGCCAT
DB	843	CGAAGCTCAC	843	CGAAGCTCAC
QY	621	TCGGAGCTGT	621	TCGGAGCTGT
DB	903	GAGGACCCAG	903	GAGGACCCAG
QY	681	GGTCAAGGCT	681	GGTCAAGGCT
DB	963	GGTGACCGTG	963	GGTGACCGTG
QY	738	CTACCTGATT	738	CTACCTGATT
DB	1023	GTTCATGATC	1023	GTTCATGATC
QY	798	GCTAAGAGAG	798	GCTAAGAGAG
DB	1083	TATGAAGGAG	1083	TATGAAGGAG
QY	858	GCTGAATCAG	858	GCTGAATCAG
DB	1143	CCTGACGCA	1143	CCTGACGCA
QY	912	ATTGACGACT	912	ATTGACGACT
DB	1203	CCTGAAGGAC	1203	CCTGAAGGAC
QY	972	GAGACGCTTA	972	GAGACGCTTA
DB	1263	GAGCACTGCT	1263	GAGCACTGCT
QY	1032	GATCAATTG	1032	GATCAATTG
DB	1323	TAAGACTAG	1323	TAAGACTAG

RESULT 14
 US-09-883-096-6
 ; Sequence 6, Application US/09883096
 ; Patent No. US20020110883A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Craven, Andrew

; APPLICANT: Yu, Ming
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Patel, Umesh A.
 ; APPLICANT: Davies, Katherine A.
 ; TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE
 ; FILE REFERENCE: 020552-001410US
 ; CURRENT APPLICATION NUMBER: US/09/883,096
 ; CURRENT FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: US 09/594,655
 ; PRIOR FILING DATE: 2000-06-15
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 1152
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: HsKip3a
 ; OTHER INFORMATION: fragment
 ; OTHER INFORMATION: Nucleotide sequence of HsKip3a fragment used in
 ; OTHER INFORMATION: ATPase assay (Figure 5).
 US-09-883-096-6

Query Match 13.1%; Score 136.6; DB 9; Length 1152;
 Best Local Similarity 51.2%; Pred. No. 3.2e-33;
 Matches 437; Conservative 0; Mismatches 389; Indels 27; Gaps 4;

QY	196	TTTGATGCTT	196	TTTGATGCTT
DB	214	TTTGACCGGT	214	TTTGACCGGT
QY	256	CCCATCTAAG	256	CCCATCTAAG
DB	274	AGGCTCTGGA	274	AGGCTCTGGA
QY	316	GGAGGTGGAA	316	GGAGGTGGAA
DB	334	GGGGCTGGGA	334	GGGGCTGGGA
QY	376	GCTCTCATGA	376	GCTCTCATGA
DB	390	-----CCT	390	-----CCT
QY	436	CTTTCGTGCA	436	CTTTCGTGCA
DB	445	TTGAGGTGCT	445	TTGAGGTGCT
QY	496	CTGTCTTCGG	496	CTGTCTTCGG
DB	505	C---CCAAGG	505	C---CCAAGG
QY	556	CTCTCCAGAA	556	CTCTCCAGAA
DB	562	CTTTCCTTCC	562	CTTTCCTTCC
QY	616	CGAAATTCGA	616	CGAAATTCGA
DB	622	CGTAACCGCA	622	CGTAACCGCA
QY	676	CTCCTGGTCA	676	CTCCTGGTCA
DB	682	TTCCAGATCT	682	TTCCAGATCT
QY	730	GGAAACTACT	730	GGAAACTACT
DB	742	GCCAAAGATG	742	GCCAAAGATG
QY	790	GGCTTCGGCT	790	GGCTTCGGCT
DB	802	GGGAGCGGCT	802	GGGAGCGGCT
QY	850	GTAGATGCGC	850	GTAGATGCGC

Db 862 CTCATGCTTGGCGGATGCGAAGGCGCGAAGACCCATGTGCGCTTACCGGACAGCAA 921
 QY 901 CTCACCTCGCTATTGACGACTCTCTGGGTGGCTCAGCCCAAGATATCTTATTGCGCAAC 960
 Db 922 CTGACCCGCGCTGCTCAAGACTCCCTCGGGGCAACTGCGGCACAGTATGATCGGTGCC 981
 QY 961 ATTGCCCTCTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGG 1020
 Db 982 ATCAGCCCTCCAGCTGACCTACGAGGACAGTACACACCCCTCAATATGCGGACCGG 1041
 QY 1021 TCCAAGGAGGTGA 1033
 Db 1042 GCCAAGGAGATCA 1054

RESULT 15

US-09-883-096-1
 ; Sequence 1, Application US/09883096
 ; Patent No. US20020110883A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Craven, Andrew
 ; APPLICANT: Yu, Ming
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Patel, Umesh A.
 ; APPLICANT: Davies, Katherine A.
 ; TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE
 ; FILE REFERENCE: 020552-00141005
 ; CURRENT APPLICATION NUMBER: US/09/883,096
 ; CURRENT FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: US 09/594,655
 ; PRIOR FILING DATE: 2000-06-15
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 4108
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Nucleic acid sequence of human kinesin motor
 ; OTHER INFORMATION: protein gene HsKip3a (Figure 1).
 ; OTHER INFORMATION: Description of Artificial Sequence: HsKip3a gene.
 US-09-883-096-1

Query Match 13.1%; Score 136.6; DB 9; Length 4108;
 Best Local Similarity 51.2%; Pred. No. 4.7e-33;
 Matches 437; Conservative 0; Mismatches 389; Indels 27; Gaps 4;
 QY 196 TTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCGAGTTTCAGTGCAG 255
 Db 344 TTTGACCGGGTCTTTGGCGAGGCGGCCACCCACAGGAGCTTCCAGCACACACGAC 403
 QY 256 CCATCTTAAGGACTTGTGGAAGGCGAGATGCGAGTGTGCTTGCCTATGAGCCACA 315
 Db 404 AGCGTCTGGACAGCTTCTCCAGGGCTACAACTGCTCAGTGTGTCCTACGGGCGCAC 463
 QY 316 GGAGCTGGGAAGACACACACATGCTGGGAGCCCGGAGCACTGGGGTATCCCGGG 375
 Db 464 GGGGCTGGGAAGACACACACATGCTGGGAAGGGAGGGGAGCCCGGCATCATGTA---- 519
 QY 376 GCTCTCATGGACCTCTGCGAGCTCACAGGAGGAGGGGTGCGAGGGCGGCGCATGGGCC 435
 Db 520 -----CCTGACCAACCGTGGAACTGTACAGGCGCTGGAGGCCCGCCAGCAGGAGAGCAC 574
 QY 436 CTTTCTGTCAACATGCTTTACCTAGATCTACAGGAGAGGATATTAGACTCTCTGAC 495
 Db 575 TTCGAGGTGCTCATCAGCTTACGAGAGGTGTATATGAACAGATCCATGACCTCTCTGGAG 634
 QY 496 CTTGCTCGGAGACTCGTATCCGAGAGACTCGCGGGGAATATCTCTGATTCGGGT 555
 Db 635 C---CCAGGGGGCCCTTGGCCATCCCGAGGAGCCCGACAGGGGGGTGGTGTCAAGGA 691
 QY 556 CTCTCCAGAGCCCATCAGTAGCTTTGTGATTTTGGAGCGGCACTTCTCTGCCAGCCAGT 615

Db 682 CTTTCTTTCCACCAGCCAGCTCAGCCGAGCAGCTGTGGAGATACTGACACAGGGGGAC 751
 QY 616 CGAAATCGGACTGTAGGAGCCACCGGCTCAACCAAGCGCTCTCCCGCAGTCACTGTGTG 675
 Db 752 CGTAACCGCAGCAGCAGCAGCCCACTGATGCCAAGGCGACTTCTCTCCGCTCCCATGCCATC 811
 QY 676 CTCTGTGTCAAGGTGACAGCGGGAACGTTTGGCCCCATTTTCGCCAG-----CGAGAG 729
 Db 812 TTCCAGATCTTTGTGAAGCAGCAGGACCGGGTTCCAGGACTGACCCAGGCTGTCTCCAGGTG 871
 QY 730 GGAATACTCTACCTGATTGACTTTGGCTGGGTGAGAGGACAAACCGGCGCACAGGCAACAG 789
 Db 872 GCCAAGATGAGCTGATTGACTTGGCTGGCTCAGAGCGGGCATCCAGACACCCATCCGAG 931
 QY 790 GGCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTTGGGCAAGTG 849
 Db 932 GGGGAGCGGCTGCGGGAGGGGGCCAAACATCAACCGCTCTCTGCTGGCGCTCATCAACGTC 991
 QY 850 GTAGATCGCTGAATCAGG-----GCTCCCTCTGTGTACCTTATCGGGACAGCAAG 900
 Db 992 CTCATATGCTTTGGCGGATGCAAGGGCGCGCAAGACCCCATGTGCCCTACCGGACAGCAA 1051
 QY 901 CTCACTCGCTTATTGACGAGCTCTCTGGGTGGCTCAGCCCAAGTATCTTATTGCGCAAC 960
 Db 1052 CTGACCCGCTGCTCAAGACTTCCCTCGGGGGCAACTCGCGCACAGTATGATGCTGCTGCC 1111
 QY 961 ATTGCCCCGTGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGG 1020
 Db 1112 ATCAGCCCTCCAGCCTGACCTACGAGGACACGTAACACACCCCTCAAAATATGCCGACCGG 1171
 QY 1021 TCCAAGGAGGTGA 1033
 Db 1172 GCCAAGGAGATCA 1184

Search completed: November 11, 2004, 01:20:41
 Job time : 558.617 secs

Result No.	Query	Score	Match			DB	ID	Description
			Match	Length				
1	1538	100.0	1538	6	AR210053	Sequence		
2	1538	99.5	1998	9	BT007259	Homo sapi		
3	1530	99.5	1998	12	BT007888	Synthetic		
4	1530	99.5	2097	6	AR304057	Sequence		
5	1530	99.5	2097	9	AB017430	Homo sapi		
6	1530	99.5	2117	9	BC028155	Homo sapi		
7	1530	99.5	2134	9	BC004352	Homo sapi		
8	1461	95.0	1464	6	AR210055	Sequence		
9	1112.4	72.3	1115	6	AR210052	Sequence		
10	1074.8	69.9	2086	10	BC003427	Mus muscu		
11	1035.4	67.3	1041	6	AR210054	Sequence		
12	885.2	57.6	151041	2	AC101919	Mus muscu		
C 13	885.2	57.6	164759	10	AC101752	Mus muscu		
C 14	744	48.4	297639	2	AC101908	Mus muscu		
15	706.8	46.0	196674	10	AC102127	Mus muscu		
16	511.6	33.3	2756	5	BC063896	Xenopus t		
17	510.2	33.2	2583	5	XLA249841	Xenopus t		
18	508.6	33.1	2565	5	BC043733	Xenopus l		
19	508.6	33.1	2640	5	BC073177	Xenopus l		

Query Match	100.0%	Score 1538;	DB 6;	Length 1538;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1538;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGCAGCGCGGGCGCTCGACGCACAGACGAGCGACGCGAGATGCGCGCAGCTTCAGCGG	60	
Db	1	ATGCAGCGCGGGCGGCTCGACGCACGAGCGACGCGAGATGCGCGCAGCTTCAGCGG	60	
Qy	61	CGCGCATCTCAGGAGCTGCTCGCTGAAGCAAGATTGCGAGTCTCTCGTCGTCCTCAC	120	
Db	61	CGCGCATCTCAGGAGCTGCTCGCTGAAGCAAGATTGCGAGTCTCTCGTCGTCCTCAC	120	
Qy	121	CTCCAGCTCGGTAAAGGTGGCTGTGCGACTCGGGCCATTTGTGATGGAACAGCGGGAG	180	
Db	121	CTCCAGCTCGGTAAAGGTGGCTGTGCGACTCGGGCCATTTGTGATGGAACAGCGGGAG	180	
Qy	181	CAAGTGATCCCCCTGTGTGCGGGGCATGGACAGCTGCTCTACAGATTGCTAACTGGA	240	
Db	181	CAAGTGATCCCCCTGTGTGCGGGGCATGGACAGCTGCTCTACAGATTGCTAACTGGA	240	
Qy	241	GGAAACCAACGAGAGACTTCAAATACCAAGTTTGATGCCCTTCTATGGGGAGAGGACTCTC	300	

Db 241 GGAACACACAGGAGCTCTCAATACCACTTTGATGCTTCTATGGGAGGAGTACTC 300
Qy 301 AGCAGGACATCTATGCAAGTTCACTGACAGCCCATCTTAAGCACTTCTGTTGAAGGGCAGA 360
Db 301 AGCAGGACATCTATGCAAGTTCACTGACAGCCCATCTTAAGCACTTCTGTTGAAGGGCAGA 360
Qy 361 ATGCCAGTGTCTTGGCTATGAGCCCAAGAGCTGGGAGAGCCACACATGCTGGGCA 420
Db 361 ATGCCAGTGTCTTGGCTATGAGCCCAAGAGCTGGGAGAGCCACACATGCTGGGCA 420
Qy 421 GCCCAGACCACTTGGGGTATCCCGGGGCTCTCATGGACCTCTGAGCTCACAAAGGG 480
Db 421 GCCCAGACCACTTGGGGTATCCCGGGGCTCTCATGGACCTCTGAGCTCACAAAGGG 480
Qy 481 AGAGGGTGGCAGAGGCGGCGCATGGGCGCTTCTGTCAACATGCTTACCTAGAGATCT 540
Db 481 AGAGGGTGGCAGAGGCGGCGCATGGGCGCTTCTGTCAACATGCTTACCTAGAGATCT 540
Qy 541 ACCAGGAGAGGTATTAGACCTCTGACCTCTGCGGAGACCTGCTAATCCGAGAAG 600
Db 541 ACCAGGAGAGGTATTAGACCTCTGACCTCTGCGGAGACCTGCTAATCCGAGAAG 600
Qy 601 ACTGCCGGGGGAATATCTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTG 660
Db 601 ACTGCCGGGGGAATATCTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTG 660
Qy 661 ATTTTGAGGCGCATCTCTGACGCAAGTGAATCGGACTGTAGGAGCCACCCGGCTCA 720
Db 661 ATTTTGAGGCGCATCTCTGACGCAAGTGAATCGGACTGTAGGAGCCACCCGGCTCA 720
Qy 721 ACCAGCGCTCTCCCGCAGTATCTGCTGCTCTGCTCAAGTGGACAGCGGGAAGTT 780
Db 721 ACCAGCGCTCTCCCGCAGTATCTGCTGCTCTGCTCAAGTGGACAGCGGGAAGTT 780
Qy 781 TGGCCCCATTTGCGCAGGAGGAGAACTCTACCTGATTGACTTGGCTGGGTCAGAGG 840
Db 781 TGGCCCCATTTGCGCAGGAGGAGAACTCTACCTGATTGACTTGGCTGGGTCAGAGG 840
Qy 841 ACAACGGCGCAGAGGCAACAGGGGCTTGGCTTAAAGAGAGTGGAGCCATCAACCT 900
Db 841 ACAACGGCGCAGAGGCAACAGGGGCTTGGCTTAAAGAGAGTGGAGCCATCAACCT 900
Qy 901 CCTGTTGCTCTGGGCAAGTGTAGATGGCTGTAATCAGGGCTCCCTGCTGCTACTT 960
Db 901 CCTGTTGCTCTGGGCAAGTGTAGATGGCTGTAATCAGGGCTCCCTGCTGCTACTT 960
Qy 961 ATCGGGAAGCAAGCTCACTGCTCTATGCAAGGACTCTCTGGGTGCTAGCCACAGTA 1020
Db 961 ATCGGGAAGCAAGCTCACTGCTCTATGCAAGGACTCTCTGGGTGCTAGCCACAGTA 1020
Qy 1021 TCCTTATGCGCAATTTGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCA 1080
Db 1021 TCCTTATGCGCAATTTGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCA 1080
Qy 1081 ACTTTGCTGCGCAGGTCAGAGAGGTGATCAATCGGCTTTTACCAATGAGAGCTCGAGC 1140
Db 1081 ACTTTGCTGCGCAGGTCAGAGAGGTGATCAATCGGCTTTTACCAATGAGAGCTCGAGC 1140
Qy 1141 CTCATGCTTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGCAA 1200
Db 1141 CTCATGCTTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGCAA 1200
Qy 1201 AGAGAGCCGAGGCGCTGAGGAAGAGAGATTTGGAGCCCTGAGCCCATGGCAGCTTCAG 1260
Db 1201 AGAGAGCCGAGGCGCTGAGGAAGAGAGATTTGGAGCCCTGAGCCCATGGCAGCTTCAG 1260
Qy 1261 CCTCTGCTCCAGAACTCAGCCCTTACAGAGAGCTTAAGCAGATGAGCCCGGCAATGC 1320
Db 1261 CCTCTGCTCCAGAACTCAGCCCTTACAGAGAGCTTAAGCAGATGAGCCCGGCAATGC 1320
Qy 1321 TGGAGCGCTCTCTCAGCTTGGACCGTCTGCTTGGCTCCACAGGGAGCCAGGGGCGCCCTC 1380
Db 1321 TGGAGCGCTCTCTCAGCTTGGACCGTCTGCTTGGCTCCACAGGGAGCCAGGGGCGCCCTC 1380

Qy 1381 TTTTGTAGTACCCAAAGCAGAGCGGATGCTGCTAATGAGACAGTAGAGAGGAGACC 1440
Db 1381 TTTTGTAGTACCCAAAGCAGAGCGGATGCTGCTAATGAGACAGTAGAGAGGAGACC 1440
Qy 1441 TAGAGATTGAGAGCTTAAAGCAAGCAAGCAAAAGAACTGGAGGGCCCAAGATGTTGCCCCAGA 1500
Db 1441 TAGAGATTGAGAGCTTAAAGCAAGCAAGCAAAAGAACTGGAGGGCCCAAGATGTTGCCCCAGA 1500
Qy 1501 AGCTGAGGAAAAAGGAGAACCATTTGTCCTCCCAATGTGA 1538
Db 1501 AGCTGAGGAAAAAGGAGAACCATTTGTCCTCCCAATGTGA 1538

RESULT 2

BT007259

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

1998 bp mRNA linear
Homo sapiens kinesin-like 4 mRNA, complete cds. PRI 13-MAY-2003

BT007259
Homo sapiens
BT007259.1 GI:30583356
FLI_CDNA.

Homo sapiens (human)
Homo sapiens
Homo sapiens

1 (bases 1 to 1998)
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Kalnina, N., Chen, X., Rolf, A., Halleck, A., Hines, L., Eisenstein, S., Koundinya, M., Raphael, J., Moreira, D., Kelley, T., Labaer, J., Lin, Y., Phelan, M., and Farmer, A.

Cloning of human full-length CDSs in BD Creator(TM) System Donor vector

Unpublished
2 (bases 1 to 1998)

Kalnina, N., Chen, X., Rolf, A., Halleck, A., Hines, L., Eisenstein, S., Koundinya, M., Raphael, J., Moreira, D., Kelley, T., Labaer, J., Lin, Y., Phelan, M., and Farmer, A.

Direct Submission
Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303, USA

This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two forms: with and without stop-codon (to allow fusion with C-terminal tag). The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the SalI and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'ACC' after SalI site and before 'ATG' to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame.

Clone distribution: <http://bioinfo.clontech.com/orfclones>

Location/Qualifiers
1. 1998
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GI:30583357"
/clone_lib="BD Creator(TM) CDS Library derived from MGC collection"
/lab_host="DH5alpha T1 resistant"
/note="Vector: pDNR-Dual"

1. 1998
/codon_start=1
/product="kinesin-like 4"
/protein_id="AAP35923.1"
/db_xref="GI:30583357"

linear PAT 12-JUN-2003

QY	786	CCATTTCCGACGAGAGGAACTCTACCTGATTGCTGGCTGGGTGAGGACCAAC	845
DB	808	CCATTTCCGACGAGAGGAACTCTACCTGATTGCTGGGTGAGGACCAAC	867
QY	846	CGGCGCAGGCAACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTG	905
DB	868	CGGCGCAGGCAACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTG	927
QY	906	TTTGTCTCTGGGCAAAAGTGGTAGATCGCTCAATCAGGGCTCTCTGCTGTGACCTTATCGG	965
DB	928	TTTGTCTCTGGGCAAAAGTGGTAGATCGCTCAATCAGGGCTCTCTGCTGTGACCTTATCGG	987
QY	966	GACAGCAAGCTCACTCGCTTATTCAGGACTCTCTGCTGGTGGCTCAGCCACAGTATCCCTT	102
DB	988	GACAGCAAGCTCACTCGCTTATTCAGGACTCTCTGCTGGTGGCTCAGCCACAGTATCCCTT	104
QY	1026	ATTGCCAAATGTCCTTCTGAGAGAGCTTCTACCTAGACACAGTCTCCGACCTCAACTTTT	108
DB	1048	ATTGCCAAATGTCCTTCTGAGAGAGCTTCTACCTAGACACAGTCTCCGACCTCAACTTTT	110
QY	1086	GCTGCCAGGCTCAAGGAGGTGATCAATCGGCTTTTACCAATCAGAGCCTGCGACCTCAT	114
DB	1108	GCTGCCAGGCTCAAGGAGGTGATCAATCGGCTTTTACCAATCAGAGCCTGCGACCTCAT	116
QY	1146	GCCTTGGGACCTGTTAAGCTGTCTCAGAAAGATTTGCTTGGTCCACAGAGGCAAGAGA	120
DB	1168	GCCTTGGGACCTGTTAAGCTGTCTCAGAAAGATTTGCTTGGTCCACAGAGGCAAGAGA	122
QY	1206	GCCTGCCGCTGAGGAGAGAGTGGAGCCCTTGGAGCCCATAGGAGCTCCAGCCTCT	126
DB	1228	GCCTGCCGCTGAGGAGAGAGTGGAGCCCTTGGAGCCCATAGGAGCTCCAGCCTCT	128
QY	1266	GCCTGCCGAGAACTCAGCCCTTACAGAGCTAAGCAGAGTGAACCGGCCATGCTGGAG	132
DB	1288	GCCTGCCGAGAACTCAGCCCTTACAGAGCTAAGCAGAGTGAACCGGCCATGCTGGAG	134
QY	1326	CGCTCTCTCAGCTTGGACCTGTCTGCTTCCCTCCAGGGAGCCAGGGGGCCCCCTCTGTTG	138
DB	1348	CGCTCTCTCAGCTTGGACCTGTCTGCTTCCCTCCAGGGAGCCAGGGGGCCCCCTCTGTTG	140
QY	1386	AGTACCCCAAGGAGAGGAGTGGTGTCTTAAGACAGTGAAGAGAGAGGACCTTAGAG	144
DB	1408	AGTACCCCAAGGAGAGGAGTGGTGTCTTAAGACAGTGAAGAGAGAGGACCTTAGAG	146
QY	1446	ATTGAGAGGCTTAAAGCAGAGCAAGAAAGAACTGAGGCCCAAGATGTTGCCAGAGGCT	150
DB	1468	ATTGAGAGGCTTAAAGCAGAGCAAGAAAGAACTGAGGCCCAAGATGTTGCCAGAGGCT	152
QY	1506	GAGGAAAGGAGAACCAATTGTCTCCCAATG	1535
DB	1528	GAGGAAAGGAGAACCAATTGTCTCCCAATG	1557
RESULT 5			
AB017430			
LOCUS			
DEFINITION			
AB017430			
ACCSSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			

Db 489 GGTGCCGAGGCGCGCCCTTCTGTCAACATGCTTTACCTAGAGATCTACCAG 548
QY 546 GAGAAAGTATTAGACCTCTCGGACCTGCTTCGGAGACCTGTAATCCGAGAGACTGC 605
Db 549 GAGAAAGTATTAGACCTCTCGGACCTGCTTCGGAGACCTGTAATCCGAGAGACTGC 608
QY 606 CGGGGGAATATCTGATTTCGGGCTCTCCAGAGGCCATCAGTAGCTTTGCTGATTTT 665
Db 609 CGGGGGAATATCTGATTTCGGGCTCTCCAGAGGCCATCAGTAGCTTTGCTGATTTT 668
QY 666 GAGCGGCACCTCTCGCAGCCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACCAG 725
Db 669 GAGCGGCACCTCTCGCAGCCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACCAG 728
QY 726 CGCTCTCCCGAGTCATGCTGTGCTCTGCTCAAGGTGACACAGGGGAACTTTGGCC 785
Db 729 CGCTCTCCCGAGTCATGCTGTGCTCTGCTCAAGGTGACACAGGGGAACTTTGGCC 788
QY 786 CCATTTCCGCGAGGAGGAAACTCTACTGATTGACTTGGCTGGGTGAGAGGACAAC 845
Db 789 CCATTTCCGCGAGGAGGAAACTCTACTGATTGACTTGGCTGGGTGAGAGGACAAC 848
QY 846 CGGCGCACAGGCAACAGGSCCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCCCTG 905
Db 849 CGGCGCACAGGCAACAGGSCCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCCCTG 908
QY 906 TTTGTCTCGGGAAGTGTAGATGCGTGAATCAGGGCCCTCCCTGTGTACCTTATCGG 965
Db 909 TTTGTCTCGGGAAGTGTAGATGCGTGAATCAGGGCCCTCCCTGTGTACCTTATCGG 968
QY 966 GACAGCAAGCTCACTCGCTATTGACGAGCTCTCTGGTGGCTCAGGCCACAGTATCCTT 1025
Db 969 GACAGCAAGCTCACTCGCTATTGACGAGCTCTCTGGTGGCTCAGGCCACAGTATCCTT 1028
QY 1026 ATTGCCAACAATGCGCTGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT 1085
Db 1029 ATTGCCAACAATGCGCTGAGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT 1088
QY 1086 GCTGCCAGGTCCAGGAGGTGATCAATCGCCCTTTTACCAATGAGAGCCTGACGCTCAT 1145
Db 1089 GCTGCCAGGTCCAGGAGGTGATCAATCGCCCTTTTACCAATGAGAGCCTGACGCTCAT 1148
QY 1146 GCCTTGGGACCTGTTAAGCTGTCTCAGAAAGATTGCTTGGTCCACAGAGGCAAGAGA 1205
Db 1149 GCCTTGGGACCTGTTAAGCTGTCTCAGAAAGATTGCTTGGTCCACAGAGGCAAGAGA 1208
QY 1206 GCCCGAGGCCCTGAGGAAGAGAGATTGGAGCCCTGAGCCCATGGCAGCTCCAGCCTCT 1265
Db 1209 GCCCGAGGCCCTGAGGAAGAGAGATTGGAGCCCTGAGCCCATGGCAGCTCCAGCCTCT 1268
QY 1266 GCCTCCAGAACTCAGCCCTTACAGAGCTAAGCGAGTGAACCGGCGCATGCTGGAG 1325
Db 1269 GCCTCCAGAACTCAGCCCTTACAGAGCTAAGCGAGTGAACCGGCGCATGCTGGAG 1328
QY 1326 CGCTCTCAGCTTGGACCTGCTTGTCTCCAGGGGAGCAGGGGGCCCTCTGTGTTG 1385
Db 1329 CGCTCTCAGCTTGGACCTGCTTGTCTCCAGGGGAGCAGGGGGCCCTCTGTGTTG 1388
QY 1386 AGTACCCCAAGGAGAGCGGATGTTGCTAATGAAGACAGTAGAGAGAGGACCTAGAG 1445
Db 1389 AGTACCCCAAGGAGAGCGGATGTTGCTAATGAAGACAGTAGAGAGAGGACCTAGAG 1448
QY 1446 ATTGAGAGGCTTAAGACGAAGCAAAAGAACTGAGGCCCAAGATGTTGGCCCAAGAGGCT 1505
Db 1449 ATTGAGAGGCTTAAGACGAAGCAAAAGAACTGAGGCCCAAGATGTTGGCCCAAGAGGCT 1508
QY 1506 GAGGAAAAGGAGAACCATTTGTCACCAATG 1535
Db 1509 GAGGAAAAGGAGAACCATTTGTCACCAATG 1538

RESULT 7
BC004352

LOCUS BC004352 2134 bp mRNA linear PRI 29-JUN-2004
DEFINITION Homo sapiens kinesin family member 22, mRNA (cdna clone MGC:1573
IMAGE:3535435), complete cds.
ACCESSION BC004352
VERSION BC004352.1 GI:13279307
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2134)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschuler, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Spletten, M., Soares, M.B., Donald, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E.,
Scherer, A., Schein, J.E., Jones, S.J. and Marra, M.A.
GENERATION AND INITIAL ANALYSIS OF MORE THAN 15,000 FULL-LENGTH
HUMAN AND MOUSE cDNA SEQUENCES
PROC. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2134)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT CONTACT: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
Parvaneh Saeedi, JR Santos, Angeliq Schnerch, Ursula Skalska,
Duane Smalios, Jeff Stott, Miranda Tsai, George Yang, Jacquie
Schein, Asim Siddiqui, Rob Holt, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 8 Row: f Column: 1
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6453817.
Location/Qualifiers
1. .2134
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:1573 IMAGE:3535435"
/tissue_type="Lung, small cell carcinoma"
/clone_lib="NIH MGC.7"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

```

gene      1. .2134
/ gene="KIF22"
/ note="synonyms: KID, OBP, OBP-1, OBP-2, KNSL4"
/ db_xref="LocustID:3835"
/ db_xref="MIM:603213"
CDS       20. .2017
/ gene="KIF22"
/ codon_start=1
/ product="kinesin family member 22"
/ protein_id="AAH04352.1"
/ db_xref="GI:13279308"
/ db_xref="LocustID:3835"
/ db_xref="MIM:603213"
/ translation="WAAGSTQQRREMAAASAAISGAGRCRLSKIGATRRPPARV
RVAVRLRFVDGTAGASPPCVKRGMDSCSLEIANRNHOETLKYOFDAFYGRSPFOOD
IYAGSVQPTIRHLHQNASVLAYGTGAGKTHWLGSPQGVIPRALMDLLOLTRE
EGAGRPWALSVTMSYLEYOEKVLDLLOPASGDLVIREDCGNLILPLSQKPLSSP
ADPERHFLPASNRVTGATRLNQRSRSHAVLLVKVDQERLAPFRQREGKYLIDLA
GSEDRNRGNGLRUKESGAINTSIFVLGVKVDALNOGLRPVFPDSDSKLRLLODSLQ
CSAHSILIANIAPERRFYLDVTALNFAARSKEVINRPTNESLQPHALGPVKLSQKE
LLGPPKARAGPBEIEGSPPEMAAPASOKLSPLQKLSMDPAMLERLISLDRLI
ASQSGAPLSTPKRERMLVMTVEKDLERLTKQKLEAKMLAQKAEKENHC
PTMLRPLSHRTVTGAKPLKXAVVMPLOLIOEQASPNAEIHLKNGKRXKLESIDL
PPEEAEDCWELQISPELLAHGRQKILDLINEGSARDLSLQRIQPKAQLIVGREL
HGFFSQVEDLERVEITGQKQMESFLKANILGLAQRCSAS"

ORIGIN
Query Match      99.5%; Score 1530; DB 9; Length 2134;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6  GCCCGGGCGGCTCGACGAGCAGAGCGGACCGGAGATGGCGGCGAGCTTCAGCGCGGCG 65
DB 23  GCCCGGGCGGCTCGACGAGCAGAGCGGACCGGAGATGGCGGCGAGCTTCAGCGCGGCG 82

QY 66  ATCTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 125
DB 83  ATCTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 142

QY 126  GCTCCGCTAAGGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 185
DB 143  GCTCCGCTAAGGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202

QY 186  GATCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 245
DB 203  GATCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 262

QY 246  CACGAGGAGACTCTCAATATACAGTTTGTATGCTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 305
DB 263  CACGAGGAGACTCTCAATATACAGTTTGTATGCTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 322

QY 306  GACATCTATGAGGTTTCAGTGCAGCCCATCTTAAAGGACCTTCTGCTGGAAGGCGAGATGCC 365
DB 323  GACATCTATGAGGTTTCAGTGCAGCCCATCTTAAAGGACCTTCTGCTGGAAGGCGAGATGCC 382

QY 366  AGTGTGCTTGCCTATGAGCCCAACAGGAGCTGGGAAGAGCACAATGCTGGGAGGCCCA 425
DB 383  AGTGTGCTTGCCTATGAGCCCAACAGGAGCTGGGAAGAGCACAATGCTGGGAGGCCCA 442

QY 426  GAGCAACTCTGGGTGATCCCGGGGCTCTCATGGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 485
DB 443  GAGCAACTCTGGGTGATCCCGGGGCTCTCATGGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 502

QY 486  GGTCCGAGGGCGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 545
DB 503  GGTCCGAGGGCGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562

QY 546  GAGAAGGTATTAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 605
DB 563  GAGAAGGTATTAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 622

QY 606  CGGGGGAATATCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTT 665

```

```

DB 623  CGGGGGAATATCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTT 682
QY 666  GAGCGGCACTTCTCGCAGCAGTCGAATCGGACTGTAGGAGCCACCCGGCTCAACCCAG 725
DB 683  GAGCGGCACTTCTCGCAGCAGTCGAATCGGACTGTAGGAGCCACCCGGCTCAACCCAG 742
QY 726  CGCTCCTCCCGCAGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 785
DB 743  CGCTCCTCCCGCAGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 802
QY 786  CCAATTTCCGACGAGAGGAGAACTCTACCTGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 845
DB 803  CCAATTTCCGACGAGAGGAGAACTCTACCTGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 862
QY 846  CGGCGCAGAGGCAACAAAGGGCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTG 905
DB 863  CGGCGCAGAGGCAACAAAGGGCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTG 922
QY 906  TTTGCTCTGGGCAAGTGGTAGATCGCTGATCAAGGGCTTCCTCTGCTGCTGCTGCTGCTGCTGCT 965
DB 923  TTTGCTCTGGGCAAGTGGTAGATCGCTGATCAAGGGCTTCCTCTGCTGCTGCTGCTGCTGCTGCT 982
QY 966  GACAGCAAGCTCAGTCGCTTATTTGAGGACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1025
DB 983  GACAGCAAGCTCAGTCGCTTATTTGAGGACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1042
QY 1026  ATTGCCAATTTGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT 1085
DB 1043  ATTGCCAATTTGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT 1102
QY 1086  GCTGCAAGTCCAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTCGAGGCTCAT 1145
DB 1103  GCTGCAAGTCCAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTCGAGGCTCAT 1162
QY 1146  GCCTTGGGACCTTTAAGCTGCTCTCAGAAAAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1205
DB 1163  GCCTTGGGACCTTTAAGCTGCTCTCAGAAAAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1222
QY 1206  GCCGAGGCGCTGAGGAGAGGAGATTTGGGAGCCTTGGAGCCCTGAGCCCATGAGGCTTCCAGCTCT 1265
DB 1223  GCCGAGGCGCTGAGGAGAGGAGATTTGGGAGCCTTGGAGCCCTGAGCCCATGAGGCTTCCAGCTCT 1282
QY 1266  GCCTCCCAAGAACTCAGCCCTTACAGAAAGCTAAGCAGCATGGACCCGCGCATGCTGGAG 1325
DB 1283  GCCTCCCAAGAACTCAGCCCTTACAGAAAGCTAAGCAGCATGGACCCGCGCATGCTGGAG 1342
QY 1326  CGCTCCTCAGCTTGGACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1385
DB 1343  CGCTCCTCAGCTTGGACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1402
QY 1386  AGTACCCCAAGCAGAGCGGATGGTGTATGAGACAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1445
DB 1403  AGTACCCCAAGCAGAGCGGATGGTGTATGAGACAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1462
QY 1446  ATTGAGAGGCTTTAAGACCAAGCAAAAGAACTGAGAGGCAAGATTTGCTGCTGCTGCTGCTGCTGCT 1505
DB 1463  ATTGAGAGGCTTTAAGACCAAGCAAAAGAACTGAGAGGCAAGATTTGCTGCTGCTGCTGCTGCTGCT 1522
QY 1506  GAGGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1535
DB 1523  GAGGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1552

```

```

RESULT 8
LOCUS      AR210055      1464 bp      DNA      linear      PAT 20-JUN-2002
DEFINITION Sequence 7 from patent US 6387644.
ACCESSION AR210055
VERSION    AR210055.1  GI:21512186
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.

```

[illegible]

QY	121	CTCCAGCTCGCTAAGGGTGGCTGTGCGACTGCGGCATTTTGTGGATGGAACAGCGGGAG	180
Db	121	CTCCAGCTCGCTAAGGGTGGCTGTGCGACTGCGGCATTTTGTGGATGGAACAGCGGGAG	180
QY	181	CAAGTATCCCCCTGTGTGGGGGATGAGAGCTGCTCTTAGAGATTGCTAACTGA	240
Db	181	CAAGTATCCCCCTGTGTGGGGGATGAGAGCTGCTCTTAGAGATTGCTAACTGA	240
QY	241	GAACCCAGCAGGAGACTCTCAATATCCAGTTTGTGCTTCTATGGGAGAGAGTACTC	300
Db	241	GAACCCAGCAGGAGACTCTCAATATCCAGTTTGTGCTTCTATGGGAGAGAGTACTC	300
QY	301	AGCAGGACATCTATGAGGTTTCAGTCCAGCCCATCTTAAGGCATTTGCTGGAAGGCGAGA	360
Db	301	AGCAGGACATCTATGAGGTTTCAGTCCAGCCCATCTTAAGGCATTTGCTGGAAGGCGAGA	360
QY	361	ATGCCAGTGTGCTTGTGCTATGAGCCACAGGAGCTGGGAAGACGACATGCTGGGCA	420
Db	361	ATGCCAGTGTGCTTGTGCTATGAGCCACAGGAGCTGGGAAGACGACATGCTGGGCA	420
QY	421	GCCAGAGCAACTGCGGGTGATCCCGGGGCTCTCATGACCTCTGAGCTCACAAGGG	480
Db	421	GCCAGAGCAACTGCGGGTGATCCCGGGGCTCTCATGACCTCTGAGCTCACAAGGG	480
QY	481	AGGAGGTCGCCAGGCGCCGCTTCTGTCTACCATGCTTCTACCTAGAGATCT	540
Db	481	AGGAGGTCGCCAGGCGCCGCTTCTGTCTACCATGCTTCTACCTAGAGATCT	540
QY	541	ACCAGAGAGGATTTAGACTCTCTGGACCTCTCGGACCTCTGAGGACCTGTAATCCGAGAAG	600
Db	541	ACCAGAGAGGATTTAGACTCTCTGGACCTCTCGGACCTCTGAGGACCTGTAATCCGAGAAG	600
QY	601	ACTGCGGGGATATCTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGCTG	660
Db	601	ACTGCGGGGATATCTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGCTG	660
QY	661	ATTTTGGCGGCACTCTCTGCGAGCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCA	720
Db	661	ATTTTGGCGGCACTCTCTGCGAGCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCA	720
QY	721	ACCAGGCTCTCTCCGAGTATGCTGTGCTCTCTGCTCAAGGTGAGCCAGCGGGAACGTT	780
Db	721	ACCAGGCTCTCTCCGAGTATGCTGTGCTCTCTGCTCAAGGTGAGCCAGCGGGAACGTT	780
QY	781	TGGCCCCATTTGCGCAGCGAGGAGAACTCTACTGATTCATCTTGGCTGGGTGAGAGG	840
Db	781	TGGCCCCATTTGCGCAGCGAGGAGAACTCTACTGATTCATCTTGGCTGGGTGAGAGG	840
QY	841	ACAACCGGCGCAGGCAACAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACT	900
Db	841	ACAACCGGCGCAGGCAACAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACT	900
QY	901	CCCTGTTTCTCTGGGCAAGTGTAGTGCCTGAATCAGGGCCCTCCCTGCTGTACCTT	960
Db	901	CCCTGTTTCTCTGGGCAAGTGTAGTGCCTGAATCAGGGCCCTCCCTGCTGTACCTT	960
QY	961	ATCGGGACAGAGCTACTCCCTATTTGAGGACTCTCTGGGTGGCTCAGGCCACAGTA	1020
Db	961	ATCGGGACAGAGCTACTCCCTATTTGAGGACTCTCTGGGTGGCTCAGGCCACAGTA	1020
QY	1021	TCCTTTATGCGCAACTTCCCTCTGAGAGCGCTTCTACTGACACAGCTCTCCGACTCA	1080
Db	1021	TCCTTTATGCGCAACTTCCCTCTGAGAGCGCTTCTACTGACACAGCTCTCCGACTCA	1080
QY	1081	ACTTGTGCGCAGGTCAGGAGGTGATCAATCG	1114
Db	1081	ACTTGTGCGCAGGTCAGGAGGTGATCAATCG	1114

RESULT 10
BC003427
LOCUS

DEFINITION	Mus musculus kinesin family member 22, mRNA (cdna clone MGC:6456 IMAGE:2615715), complete cds.		
ACCESSION	BC003427		
VERSION	BC003427.1 GI:13097359		
KEYWORDS	MGC.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	1 (bases 1 to 2086)		
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Dietzenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schuetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McSwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Faney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywicki, M.I., Skalek, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
PUBMED	12477932		
REFERENCE	2 (bases 1 to 2086)		
AUTHORS	Strausberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: Jeffrey Green M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LiNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amg@bcm.tmc.edu Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegad, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.		
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IPAK Plate: 5 Row: m Column: 7 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21704181.		
SOURCE	Location/Qualifiers 1. .2086 /organism="Mus musculus" /mol_type="mRNA" /strain="FVB/N" /db_xref="taxon:10090" /clone="MGC:6456 IMAGE:2615715" /tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating ductal carcinoma 5 month old virgin mouse." /clone_lib="NCI CGAP_Mam6" /lab_host="DH10B" /note="Vector: pCMV-SPORT6" 1. .2086 /gene="Kif22"		
gene			

/note="synonym: Kid"
 /db_xref="LocusID:110033"
 /db_xref="MGI:109233"
 20. .2002
 CDS
 /gene="Kif22"
 /codon_start=1
 /product="kinesin family member 22"
 /protein_id="AAH03427.1"
 /db_xref="GI:13097360"
 /db_xref="LocusID:110033"
 /db_xref="MGI:109233"
 /translation="MSLRKATCPORREMASATSPGRVSKGLGRPPILARVRVAVR
 LRPMDGETAKELPCVRAIDSCLELVANWKQETILKYOFDAFYGEKSKTQOEYVVS
 VQPLRLHLLGQNASVLAYGPTGAGKTHWLGSPQGVIPRALMDLLQALRESABG
 RPDVSVAMSYLEYIQKVLDDLPASGLVIREDCRGNILIFELTKPITISDFEQ
 HFLPASNRVAVGTLNORSRSHAVLLVQDQRELRTPFRQREGKLLIDLAGSEDN
 RRTGNOIRIKESGAINISFLVGLKVVVDALNOGLPRIPYRDSKLALPVLKSLQELLGSPS
 ILIANIAPFERFYDTTISALNFTARSKEVINRPNESLOPHALAPVLLKSLQELLGSPS
 EAKKAGPEESTGSPSTAPASAKSLLOKLSNMDPAMLENLISMERLLGSGQS
 QGTPLANTPKRERVMKTVEEKNLETERLKMKELEAKVLAQEAADPDPREKENTPTI
 LQPPASVSGTVKPLKAVNMPLORIQKRESSNOIQLLKKPKRKLPSPESEAVEK
 DEVDWEVQISPELLAHGRKKLLDLNLSGARELSLQRIQIGOKKAQLIVGWRELHGFPS
 EVEDLEQVEGISGRQVESFLKANLLSLAASQHSQPS"

ORIGIN

Query Match 69.9%; Score 1074.8; DB 10; Length 2086;
 Best Local Similarity 83.6%; Pred. No. 2.7e-269; Mismatches 238; Indels 0; Gaps 0;
 Matches 1214; Conservative 0;

Qy	59	GGCGCGGATCTCAGGAGCTGTGCTGCTGCGCTAAGCAAGATTGGAGCTACTCGTCGTC	118
Db	61	GGCGTGGCCACCTCAGGACCTGCGCCGCTGTGTAAGCAAGAGGGGCTTGGCCGGCGTCC	120
Qy	119	ACCTCCAGCTCGCTAAGGGTGGCTGTCGCACTCGGCCCAATTTGTGATGCAACAGCGGG	178
Db	121	GCCTCTAGCTCGAGTAAGGGTGGCTGTCGCGCTGCGCCCGTTTATGATGAGAGACAGA	180
Qy	179	AGCAAGTGATCCCGCTGTGTCGGGGCATGGACAGCTGCTCTACAGATTGCTACTG	238
Db	181	AGCGAAGGAGTCCCGCTGTGTCGAGGCATAGACAGCTGCTCTTGAAGTGGCTAACTG	240
Qy	239	GAGGAAACACAGGAGACTCTCAATACCAATACCACTTTGATGCTTATGGGAGAGAGTAC	298
Db	241	GAAGAAATACAGGAGACGCTCAATATCAGTTTGTATGCTTCTATGCGAGAGAGACAC	300
Qy	299	TCAGCAGACATCTATGAGCTTCAGTCAGCCCATCTTAAGGCATTTGCTGGAAGGGCA	358
Db	301	TCAGCAGAGAGTCTATGATAGTTTCAGTACAGCCCATCTTAAGGCATTTGCTGGAAGGGCA	360
Qy	359	GAATGCCAGTGTCTTGCCTATGACCCACAGGAGCTGGGAGACGACACCAATGCTGGG	418
Db	361	GAATGCCAGTGTCTTGCCTATGAGCCCTATCGGGGAGGGAAGACACCAATGCTGGG	420
Qy	419	CAGCCACAGACCACTGGGGTGTATCCCGGGCTCTCATGGACCTCTGCGAGCTCAACAG	478
Db	421	CAGCCACAGACCACTGGAGTGTCTTGGGGCTCTCATGGACCTCTGCAACTAGGACAG	480
Qy	479	GGAGAGGGTCCGAGGGCGGCGCATGGGCGCTTTCTGTCAACCATGTCTTACCTTAGAGAT	538
Db	481	GGAGAGAGTGTCTGAGGACGGCCATGGGACGTTTCTGTGTATGTCTTATTAGAGAT	540
Qy	539	CTACCCAGGAGAGTATTAGACCTCTGACCCCTGCTTGGGAGACCTGCTTAATCCGAGA	598
Db	541	CTACCCAGGAGAGTATTAGACCTCTTGGGCTCTGCAATCAGGAGACCTGCTGTATCCGCGA	600
Qy	599	AGACTCGGGGGGAATATCTGTATTCGGGGTCTCTCCAGAGGCCCATCAGTAGCTTTGC	658
Db	601	AGACTCGGGGGGAATATCTGTATTCGGGGTCTCTCCAGAGGCCCATCAGTAGCTTTGC	660
Qy	659	TGATTTTGGAGGGCACTTCTGCGCAGCAGTTCGAAATCGGACTGTAGGAGCCACCGGCT	718
Db	661	TGACTTCGAGCAGCACTTCTCTTCCAGCGCAGTAGAATCGAGCTGTAGGAGCCACCGGCT	720

Qy	719	CAACACGCGCTCTCCGCGAGTCATGCTGTGCTCTCGTCAAGTGACGACGAGGGGACG	778
Db	721	TAACACGCGCTCTCCGCGTAGTCACGCGAGTCTCTTGGTCAAGTAGATCAGCGTAGACG	780
Qy	779	TTTGGCCCGCATTTTCGCCAGCGAGAGGAGAACTCTACCTGATTGACTTTGGCTGGTTCAGA	838
Db	781	TTTGGACTCCATTTTCGCCAGCGGAGGAGAACTCTACCTGATTGACTTTGGCTGGTTCAGA	840
Qy	839	GGACCAACCGCGGACAGGCAACAGGGCCTTCGGCTTAAAGAGAGTGAGGAGCATCAACAC	898
Db	841	GGACCAACCGTTCGACAGGCAACAGGGCATTTCGGCTTAAAGAGAGTGAGGAGCATCAACAC	900
Qy	899	CTCCCTGTTTGTCTCGTGGCAAGTGTAGATGCGTGAATCAGGGCCTCCCTGCTGTAC	958
Db	901	CTCCCTGTTTGTCTCGTGGCAAGTGTGTGATGATTAACACAGGGCCTCCCTGCTGTAC	960
Qy	959	TTATCGGACAGCAAGCTCACTCGCCCTATTTCGAGGACTCTCTCGGTGGCTCAGGCCACAG	1018
Db	961	ATACCGGACAGCAAGCTCACTCGTCTGCTGCGAGGACTCTCTCGGAGGCTCAGCTCATAG	1020
Qy	1019	TATCCTTATTGCCAACATTGCCCTCTGAGAGACGCTTCTACTAGACACAGTCTCCGCACT	1078
Db	1021	CATCTCTATTGCCAACATTGCTCTCGAAGACGCTTTTACGAGATACAAATTCAGCATT	1080
Qy	1079	CAACTTTGCTGCCAGGTCCTCAAGGAGTGTATCAATCGGCCCTTTTACCAATCAGAGCCCTGCA	1138
Db	1081	GAACCTTCACTGCTGAGTTCAGGAGGAGTGTATTAACCGGCCCTTTTACCAATGAGAGCTTACA	1140
Qy	1139	GCCTCATGCTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTTGTTGGTCCACACAGGCG	1198
Db	1141	GCCTCATGCTTGGCACCCTGTTAAGCTGTCTCAGAAAGAAATTTGTTGGTCCACAGGCG	1200
Qy	1199	AAAGAGAGCCCGAGGCGCTCAGGAAAGAGAGATTTGGAGCCCTTGAGCCCATGGCAGCTCC	1258
Db	1201	AAAGAGAGCCCGAGGCGCTCAGGAAAGAAATCAACTGGGAGCCCTGAGCTTACAGCAGCTCC	1260
Qy	1259	AGCTCTGCTCCCGACAACTCAGCCCTCAGAAAGAAATTTGTTGGTCCACACGCGGCAT	1318
Db	1261	AGCTCTGCTCCCGACAACTAAGCCCTTTCAGAAAGAAATTTGTTGGTCCACGCGCAT	1320
Qy	1319	GCTGGAGCGCTCTCTCAGCTTGGACCGTCTGCTTGGCTTCCAGGGGAGCCAGGGGGCCCC	1378
Db	1321	GTTGGAGAACCTCTCTGAGTATGAAAGTCTGCTGGTTCCTCCAGGAGCCACAGGACCCC	1380
Qy	1379	TCGTGTGAGTACCCAAAGAGAGCGGATGGTGTCTAAATGAAGACAGTGAAGAGAGGA	1438
Db	1381	CCTACTGAATACCCAAAGAGAGCGAATGGTGTCTCATGAAGACAGTGGAGGAGAAAGAA	1440
Qy	1439	CCTAGAGATTGAGAGGCTTAAGACGAGCAAGCAAAAGAACTGGAGGCCAAGATTTGGCCCA	1498
Db	1441	CTTAGAGATTGAGAGGCTTAAGATGAAGCAAAAGAACTGGAGGCCAAGATTTGGCCCA	1500
Qy	1499	GAAGGCTGAGGA 1510	
Db	1501	GGAGGACCCGA 1512	

RESULT 11
 AR210054
 LOCUS
 DEFINITION Sequence 5 from patent US 6387644.
 ACCESSION AR210054
 VERSION AR210054.1 GI:21512185
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1041)
 AUTHORS Beraud, C.
 TITLE Motor proteins and methods for their use
 JOURNAL Patent: US 6387644-A 5 14-MAY-2002;
 FEATURES Location/Qualifiers
 source 1..1041

Submitted (14-MAR-2003) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 14, 2003 this sequence version replaced gi:28631351.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L17745
Center clone name: 158 A.6
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 150558 bases at least Q40
Consensus quality: 150654 bases at least Q30
Consensus quality: 150687 bases at least Q20
Insert size: 151000; agarose-fp
Insert size: 150741; sum-of-contigs
Quality coverage: 11.8 in Q20 bases; agarose-fp
Quality coverage: 11.8 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 33517: contig of 33517 bp in length
33518 33617: gap of 100 bp
33618 42736: contig of 9119 bp in length
42737 42836: gap of 100 bp
42837 84394: contig of 41558 bp in length
84395 84494: gap of 100 bp
84495 151041: contig of 66547 bp in length.
Location/Qualifiers
1. .151041
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP24-158A6"
/clone_lib="RPCI-24 Male Mouse BAC"
1. .33517
/note="assembly_fragment
clone_end:SP6
vector_side:left"
33618. .42736
/note="assembly_fragment"
42837. .84394
/note="assembly_fragment"
84495. .151041
/note="assembly_fragment
clone_end:T7
vector_side:right"

Query Match 57.68; Score 885.2; DB 2; Length 151041;
Best Local Similarity 78.28; Pred. No. 1.1e-219;
Matches 1136; Conservative 0; Mismatches 288; Indels 28; Gaps 5;
57 GCGCGCGCGATCTCAGGAGTGTGCTGCGGTAAGCAAGATGGAGCTACTCGTCTG 116
65725 GCGTCAGCCACTCAGACCGGGTCTTTGT---ATAAGCAAGGTAGGATGGCTGTCTG 65669
117 CCACCTCAGCTCCGTAAGGTGGCTGTGGCATCTCGGCCATTTGTGGATGGAACAGCG 176
65668 CTGCTTCTAGCTGAGTAAGGGTGGCTATATAATTGTGCTGTTTGTGGATGAAGACA 65609

QY 177 GGAGCAAGTATGCCCTCTGTGCGGGCATGACAGCTGTCTCTTAGAGATTGCTAAC 236
DB 65608 GAAGCGAGGAGCACCCTCTGTACAGCCATAGACGCTGTCTCTTGAAGTGGCTAAC 65549
QY 237 TGGAGGAACACACAGGAGACTCTCAAATACCACTTTGATGCCCTTCTATGGGAGAGAGT 296
DB 65548 TGGAGGAATATACAGGAGACTATCAAATATCACTTTGATGCCCTTCCATGGAGAGAGC 65489
QY 297 ACTCAGCAGGACATCTATGAGGTTGATGTCAGCCCATCTAAGGACCTTGTCTGGAAGGG 356
DB 65488 ACTCAGTGGGACATCTATGTAGGTTTAGTACAGCCCATATAAGGAATCTTGTCTGGAAGGG 65429
QY 357 CAGAAATGCCAGTGTCTTGTGCTATGACCCACAGGAGCTGGGAGAGCGACACAATCTG 416
DB 65428 CAGAAATGCCAGTGTCTTGTGCTATGACCCATGGGAGGAGAGA--CACACATGTTG 65371
QY 417 GGCAGCCCAAGAGCAACCTGGGGTGATCCCGCGGGCTCTCATGGACCTCTGCGAGCTCAC 476
DB 65370 GGCAGCCCAAGAGCAACCTGGGGTGATCCCTCAGACTCTCCAGACCTCTCTGCACCTAAC 65311
QY 477 AGGAGAGGAGTCCGAGGCGCGGATGATGGGCGCTTCTGTCAACATGTCTTACCTAGAG 536
DB 65310 AGGAGAGAGAAATCTCAGGCGCTGGCCATGGGACATTTCTGTCTTGTCTATTTAGAT 65251
QY 537 ATCTACAGGAGAGAGTATTAGACCTCTCGGACCTGTCTCGGAGACCTGTAGAGCCACCG 596
DB 65250 CTCTACAGGAGAGTATTAGACCTCTTGGATCTCGATCAGAAGACCTTAGTGTCCGA 65191
QY 597 GAAGACTGCGGGGGGAATATCTGTATTCGGGGTCTCTCCAGAGAGCCCATCAGTAGCTTT 656
DB 65190 GAAGGCTGACAGGAGAAATCTGTATTCAGAGCTCTCACAGAGAGCCCATCAGTAGTTT 65131
QY 657 GCTGATTTGAGCGGCACTCTCTGCGAGCCAGTGCAGAAATCGGACTGTAGAGCCACCG 716
DB 65130 TCTGACTTCGAGAAGCACTCTCTTCTAGTCAGTAGAATCGAGTTGTAGAGCTACTCTG 65071
QY 717 CTCACACAGCGCTCTCCGCGAGTCTGTCTCTGCTCAAGGTGGACACGCGGAA 776
DB 65070 CTTAACACGCGCTCTTGGCGGTAGTCAITGACGCGCTCTTGGTCAAGGTAGATCAGCATG 65011
QY 777 CGTTTGGCCCCATTTCCGCGAGGAGGGAATCTTACCTGATTTGACTTGGCTGGGTCA 836
DB 65010 CGTTTGTACTCTCAATTCGCGAGTGGGAGGAAATCTTACCTTATTGATTTGGCTGGTTCA 64951
QY 837 GAGCAACACC-----GCGGCAACAGCAACAGGCGCTTCGGCTAAAGAGAGTGGAGCAT 892
DB 64950 GAGGCAACCGTCAGTCGCGACAGGCAACCGAGGCAATCTCTCTCAAGAGAGTGGAGCAT 64891
QY 893 CAACACCTCCTCTGTTGTCTGCGGCAAGTGGTAGATGCGCTGAATCAGGCGCTCCCTCG 952
DB 64890 CAACACCTCCTCTTGTGTTACTGGGCAAGTAGTGGATGATTAACCAAGGCGCTCCCTCA 64831
QY 953 TGTACCTTATCGGAGCAGCAAGCTCTCTCGCTTATTCAGGAGCTCTCTGGGTGGCTCAGC 1012
DB 64830 TATACCATACACAGGACAGCAGGCTCACTGCTCTGTTGCAAGACTCTCTGAGAGGCTCAGC 64771
QY 1013 CCACAGTATCTTTATTTGCCAATGTCGCCCTTAGAGAGCGTCTTACCTAGACACAGTCTC 1072
DB 64770 TCATAGCATCTCTCATTTGTGTCAGCATTTGCTCTCTGAAAGATGTTTATCTGATCAATCTC 64711
QY 1073 CGCACTCAACTTTGCTGCCAGGTCCAGGAGGTGATCAATCGGCGCTTTTACCATAGAG 1132
DB 64710 AGCATTTAACTTCACTGTAGTTCAGAGGAGGTGATTAACAGGCTTTTACCAAGGCGAG 64651
QY 1133 CTTGACGCTCATGCGCTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACC 1192
DB 64650 TGTACAACTCAGGCGCTTGACACCTGTTAAGCTGTCTCAGAAAGATCTGCTAGTTTATC 64591
QY 1193 AGAGGCAAGAGAGCCGAGGCGCTTAGGAGAGAGATGGGAGCGCTGAGGCCCTGAGCCCTGCG 1252
DB 64590 AGAGGCAAGAGAGCCGAGGCGCTTAGGAGAGAGATCAACTGGGAGCGCTGAGTCTATAGC 64531
QY 1253 AGCTCCAGCGCTCTGCTCTCCAGAAAATCTCAGCGCCCTACAGAACTAAGCATGAGACCC 1312

ORIGIN

Query Match 57.68; Score 885.2; DB 2; Length 151041;
Best Local Similarity 78.28; Pred. No. 1.1e-219;
Matches 1136; Conservative 0; Mismatches 288; Indels 28; Gaps 5;
57 GCGCGCGCGATCTCAGGAGTGTGCTGCGGTAAGCAAGATGGAGCTACTCGTCTG 116
65725 GCGTCAGCCACTCAGACCGGGTCTTTGT---ATAAGCAAGGTAGGATGGCTGTCTG 65669
117 CCACCTCAGCTCCGTAAGGTGGCTGTGGCATCTCGGCCATTTGTGGATGGAACAGCG 176
65668 CTGCTTCTAGCTGAGTAAGGGTGGCTATATAATTGTGCTGTTTGTGGATGAAGACA 65609

Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuppback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Teskaye, S., Theodore, J., Topham, K., Travers, M., Vasiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (17-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (Bases 1 to 164759)

Birren, B., Nusbaum, C., Lander, E., Aboueleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barina, N., Bastien, V., Bloom, T., Boguslavsky, L., Boughgali, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearrellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.

Hagopian, J., Gargya, S., Graham, L., Grand-Pierre, N., Hafez, N.,
 Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
 Johnson, R., Jones, C., Kamat, A., Karacas, A., Kells, C., Landers, T.,
 Levine, R., Lindblad-Toh, K., Liu, G., Liu, A., Lui, A., Mabbitt, R.,
 MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,
 McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
 Rettung, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
 Runka, R., Rice, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Stubbs, M., Talamas, J., Tsafaye, S., Theodore, J.,
 Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R.,
 Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
 Zimmer, A. and Zody, M.

Direct Submission
 Submitted (23-APR-2004) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 23, 2004 this sequence version replaced gi:45504330.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center
 Center: Whitehead Institute/MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence.submissions@broad.mit.edu
 ----- Project Information
 Center project name: LI7344
 Center clone name: 344_C18

 Location/Qualifiers
 1. 164759
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="1"
 /map="1"
 /clone="RP24-344C18"

```

/clone=RP24-34418
/clone_lib="RPCI-24 Male Mouse BAC"
1. .1655
  /rpt family="L1_MM"
1. .4
  /note="clone boundary
  clone_end,SP6
  site:MbolI"
1767. .1787
  /rpt family="AT_rich"
2115. .2157
  /rpt family="AT_rich"
2236. .2258
  /rpt family="AT_rich"
3252. .3420
  /rpt family="Lx7"

```



```
Db 145182 TCATAGATCCTCATTTGTCAGCATGCTCTGCTGAAGATGTTTTTACCTGATCAATCTC 145123
QY 1073 CGCACTCAACTTTGCTGCGAGGTCGAAGGAGTGATCAATCGGCCCTTTTACCAATGAGAG 1132
Db 145122 AGCATTAACCTTCACTGCTAGGTCGAAGGAGTGATTAACCAAGCTTTACCAACGAG 145083
QY 1133 CTGCGAGCCTCATGCTCTGGGACCTGTTAAGCTGCTCTCAGAAAGAAATGCTTGGTCCACC 1192
Db 145062 TGTACAACTTCAGGCCTTGACACCTGTTAAGCTGCTCTCAGAAAGATCTGCTAGGTTTATC 145003
QY 1193 AGAGGCAAGAGAGCGCGGCGCTCAGGAAGAGGAGATGGAGCGCCTGAGCCCATGGC 1252
Db 145002 AGAGGCAAGAGAGCGCGGCGCTCAGGAAGAGGAGATCAACTGGGAGCGCCTGAGTCTATAGC 144943
QY 1253 AGCTCAGGCTCTGCTCTCCAGAAACTCAGCCCTTACAGAACCTTAAGCAGCATGACCC 1312
Db 144942 ATCTCAGCATCTGCTCTCCAGAA-----CTAAGCAATATGATCT 144901
QY 1313 GGCATGCTGGAGCGCTCTCTCAGCTTGGACCGTCTGCTTGCCTCCAGGGAGCCAGGG 1372
Db 144900 AGCCATGTTGGGAAAGCTCTCTGAGTATGGAATGTTTCTGGTGTCTCAAGGGAGCCAGAG 144841
QY 1373 GGCCTCTCTGAGTACCCCAAGCAGAGCGGATGCTGCTTAATGAAGACAGTAAAGA 1432
Db 144840 GAACCTCTACTGAATACCCCAA--CGAGAGCGAATGCTGCTCATTAAGACGCTGGAGGA 144782
QY 1433 GAAGGACCTTAGAGATGAGAGCGCTTAAGACGAGCAAGAAAGAACTGAGGCCAAGATGTT 1492
Db 144781 CAGATATTGGAGATTGAGAGGCTTAAGATGAAGCAAGAAAGAACTGGAAGCCAGGTGCT 144722
QY 1493 GGCCCAAGAGGC 1504
Db 144721 GGTCCAGGAGC 144710
```

```
RESULT 14
AC101908/c
LOCUS
DEFINITION Mus musculus chromosome 1 clone RP24-275J1 map 1, *** SEQUENCING IN
PROGRESS ***, 8 unordered pieces.
ACCESSION AC101908
VERSION AC101908.4 GI:41199137
KEYWORDS HTG; HTGS PHASE1; HTGS FULLTOP; HTGS_ACTIVEFIN.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 297639)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Mus musculus chromosome 1, clone RP24-275J1
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 297639)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campioano,A., Chang,J., Chazaro,B.,
Chapel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
Cook,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LeRoque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosett,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
```

```
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 297639)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Chapel,Y., Collamore,A., Cook,A., Cooke,P., Corum,B.,
DeArelano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (18-JUN-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 27, 2004 this sequence version replaced gi:31880232.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
----- Project Information
Center project name: L17732
Center clone name: 275_J_1
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 63747: contig of 63747 bp in length
* 63748 63847: gap of 100 bp
* 63848 73892: contig of 10045 bp in length
* 73892 73992: gap of 100 bp
* 73992 154450: contig of 80458 bp in length
* 154450 154550: gap of 100 bp
* 154550 167808: contig of 13258 bp in length
* 167808 167908: gap of 100 bp
* 167908 176711: contig of 8803 bp in length
* 176711 176811: gap of 100 bp
* 176811 233500: contig of 56689 bp in length
* 233500 233600: gap of 100 bp
* 233600 241222: contig of 7622 bp in length
* 241222 241322: gap of 100 bp
* 241322 297639: contig of 56317 bp in length.
* 297639 Location/Qualifiers
* 1.297639
* /organism="Mus musculus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10090"
* /chromosome="1"
* /map="1"
```

Db	250	TCATAGCATCCTCATTGTTCAGCATTTGCTCTCTGAAAGATGTTTTTACCTGGATACATCTC	191
Qy	1073	CGCACTCAACTTTTGCTGCCAGGTCCTGAGGAGGTGATCAATCGCGCTTTTACCAATGAGAG	1132
Db	190	AGCATTAACCTTTCACCTGCTAGGTCCTGAGGAGGTGATTAACACGAGCTTTACCAAAAGCGAG	131
Qy	1133	CCTGACGCTCATGCTTGGACCTGTTAAGCTGTCTCAGAAAGATGCTTGGTCCACC	1192
Db	130	TGTACACCTTCAGGCTTTGACACTGTTAAGCTGTCTCAGAAAGATGCTGCTGTTTATC	71
Qy	1193	AGAGGCAAGAGAGCGCCGCTGAGGAAGGAGATTTGGAGCGCTCAGGCCATGGC	1252
Db	70	AGAGGCAAGAGAGCGCCCTCAGGAAGATCAACTGGAGCGCTCAGTCTATAGC	11
RESULT 15			
AC102127		196674 bp	DNA
LOCUS			linear
DEFINITION			Mus musculus chromosome 1, clone RP23-202A19, complete sequence.
ACCESSION			AC102127
VERSION			AC102127.11
KEYWORDS			HTG.
SOURCE			Mus musculus (house mouse)
ORGANISM			Mus musculus
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS			1 (bases 1 to 196674)
TITLE			Barren,B., Nusbaum,C. and Lander,E.
JOURNAL			Mus musculus chromosome 1, clone RP23-202A19
REFERENCE			Unpublished
AUTHORS			2 (bases 1 to 196674)
			Barren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouckhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Katat,A., Karatas,A., Kells,C., LaRoque,K., Lanazare,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Riback,M., Rile,R., Rise,C., Rogov,P., Roman,J., Rosetti,N., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Traversman,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE			Direct Submission
JOURNAL			Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE			3 (bases 1 to 196674)
AUTHORS			Barren,B., Nusbaum,C., Lander,E., Abouelail,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Bouckhalter,B., Camarata,J., Chang,J., Bloom,T., Boguslavsky,L., Cook,A., Cooke,P., Corum,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fero,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Katat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., MacLean,C., MacDonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,

$\gamma = \frac{1}{\sqrt{1 - \beta^2}}$

Query Match 46.0%; Score 706.8; DB 10; Length 196674; Best Local Similarity 73.3%; Pred. No. 4.2e-173; Matches 1013; Conservative 0; Mismatches 357; Indels 12; Gaps 8;									
QY	143	TGTCGACTGGCGCAATTCGTGATGGAACAGCGGGAGCAAGTGATCCCCCTGTGTGG	202						
Db	101632	TTTATGACTATGCCATTTATGATGAAGAAAAGAAAGGATGCCCTTGTGTAG	101691						
QY	203	GGGATGGACAGCTGCTCTTAGAGATTGTAACTGGAGGAACACAGGAGACTCTCAA	262						
Db	101692	AGGATATAAGCTGCTCTCTTGAGGTGGCCACCTTGAGAAATCTTATGAGAATACTAA	101751						
QY	263	ATACAGTTTGTATGCCCTTCTATGGGGAGAGAGTACTCAGCAGGACATCTATCGAGTTTC	322						
Db	101752	ATATCAGTTTAAATGCTCTTTATTTAGAGAAAGAGCACTCAGCAGAAACATCTAAGTAGGTTTC	101811						
QY	323	AGTCAGCCCATCTTAAGGCACTTGTGGAAGGCGCAATGCCAGTGTGCTTGCCTATGG	382						
Db	101812	AGTACAGCTATCTGTGGCACTTGTCTGGATGGCAGATACCAAGTGTGCTTACATATGG	101871						
QY	383	ACCACAGGAGCTGGGAAGCGCACAAATGCTGGGAGGCC-----AGACAACTGGGG	438						
Db	101872	GACTATTGAGCGAGGAACACATATAATCTAGGTAGCCCATGGGAACAAATGGAG	101931						
QY	439	TGATCCCGGGCTCTCATGAGACCTCTGCAGCTCACAAGGAGGAGGTGCCGAGGCC	498						
Db	101932	TGATTCCTTAGCCCTCTCATAGACCTCTGCAGATTAACAGGGGAAGAAAGTGCTGAAGGCT	101991						
QY	499	GGCATGGGCTTCTGTACCACTTCTTACCTAGAGATCTACCAAGGAGAGGTATTAG	558						
Db	101992	GGCATGGGACATTTCTGTAGGATGTCCTATTAGAGTCTACCAAGGAGAGGTATTAG	102051						
QY	559	ACCTCTTGAACCTCTGCTTGGGAGACCTGTGTAATCCAGAGAGACTGCCGGGGGAATACC	618						
Db	102052	ACCTCTTATATCTGTCATCAGGAGACCTAGTGATATGAGAAGATTGTGAGGGAAATACC	102111						
QY	619	TGATTCGGGTCTCTCCAGAGGCCCATCAGTACTTGTGATTTTGGGCGGCACTTCC	678						
Db	102112	TGATTCAGTTCTCACCAGAGCCCAATCTTCTCTCTCTGACTTTGAGCAGTACTT-C	102170						
QY	679	TGCCAGCGAGTCGAAATCGACTAGGAGCCACCGGCTCAACAGCGCTCTCTCCCGCA	738						
Db	102171	TGCCAGCGCTGAGAAATCAAGTTGTAGGAGTACTCTCTTAAAGCAGTCTCTTACCATA	102230						
QY	739	GTCATGCTGCTCTCTGCTGAAGTGGACAGCGGGAACGTTTGGCCCCCATTTCCGCGC	798						
Db	102231	-TCATTCAATGCTCTTTATCAAGGTGGATCAGCATGGACATTTGACTCTGTTTCACTAGT	102289						
QY	799	GAGAGGGAATACTCTACTGATTGACTTGGCTGGGTGAGAGCAACCGGCGCACAGGCA	858						
Db	102290	GAAAGGHAATACTCTACCATATTGATTTGGTTGGCTCAGAGGACCAACCAACA-AGGA	102348						
QY	859	ACAAAGGGCTTCGGCTAAAGAGAGTGAGCCATCAACACCTCCCTGTTGTCTCGGCA	918						
Db	102349	ACCAGTGATTTCAACTCAAGAGAGTGGCGCATCAACACCTGCTTC-TTGTACTGGCA	102407						
QY	919	AAGTGGTAGATCGGTGAAATCAGGGCTCTCTCTGTGACTTATCGGAGCAGCAAGCTCA	978						
Db	102408	AAGTGGAGTTGCAATCATATTAGGGCTCCCTCTCTATATTACTGGGACAGAAAAATCA	102467						
QY	979	CTGGCTATTGCGAGCTCTCTGGGTGGCTCAGGCCACAGTATCTTATTGCCAACATTG	1038						
Db	102468	TTTGCTTATTGCGAGCTCTCTGGAGAGCTCAGCTCATAACATCATATTACCAACACTG	102527						
QY	1039	CCCTGAGAGAGCTTCTTACTAGACACAGTCTCGGCACTCAACTTTGCTGCCAGGTCCA	1098						
Db	102528	CTCCTGAAGACATTTTACTTGAATCATCTGAGCATTTGAATTTCACTCTAGGTCCA	102587						
QY	1099	AGGAGTGCATCAATCGGCTTTTACCAATGAGCGCTCAGCCTCATGCCCTTGGGACCTG	1158						
Db	102588	AGGAGCTGATTAATTTGGCTTTTACCAATGGAGTCTACAGCCTCATGCCCTACACCTG	102647						
QY	1159	TTAAGCTGCTCAGAAAGAAATTTGTTTGTCCACAGAGGCAAGAGAGCCCGAGGCCCTG	1218						

Search completed: November 10, 2004, 13:18:01
Job time : 6803.66 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2004, 09:45:10 ; Search time 795.24 Seconds
(without alignments)
10152.424 Million cell updates/sec

Title: US-10-797-893-3
Perfect score: 1538
Sequence: 1 atgcagccgcggcggtcgc.....accattgtccacaatgtga 1538

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 4134886 seqs, 2624710521 residues
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1538	100.0	1538	10	Adc23339 DNA encod
2	1538	100.0	1538	12	Adq60229 Human mic
3	1530	99.5	2097	12	Adq09241 Human KNS
4	1519	98.8	2099	10	Adk40995 Novel hum
5	1461	95.0	1464	10	Adc23343 DNA encod
6	1461	95.0	1464	12	Adc23343 Human mic
7	1450.2	94.3	2104	3	Aaf15853 Human pro
8	1112.4	72.3	1115	10	Adc23337 DNA encod
9	1112.4	72.3	1115	12	Adq60227 Human mic
10	1035.4	67.3	1041	10	Adc23341 DNA encod
11	1035.4	67.3	1041	12	Adq60231 Human mic
12	480	31.2	491	6	Abk70260 Human lun
13	376.4	24.5	386	4	Aas38879 Novel hum
14	265	17.2	464	9	Ach44493 Human toe
15	233.2	15.2	531	12	Ach77408 Human gen
16	230.4	15.0	232	12	Ach91108 Human gen
17	185.6	12.1	2095	11	Adm01564 Human CDN
18	185.6	12.1	3624	12	Ado44181 DNA encod
19	156.8	10.2	3374	10	Adj95073 Novel NOV
20	154.8	10.1	2375	8	Abx34596 Human mdd
21	138.6	9.0	4108	6	Abx94614 Human kin

22	138.6	9.0	4108	8	Abx57215 Human kin
23	137.2	8.9	2675	10	Adj95077 Novel NOV
24	136.6	8.9	1152	6	Abx94616 Nucleotid
25	136.6	8.9	1152	8	Abx57216 DNA encod
26	128.8	8.4	897	4	Aai70088 Human kin
27	128.8	8.4	897	6	Aai39619 DNA encod
28	128.8	8.4	897	12	Adm81051 Human HsK
29	127.8	8.3	488	5	Abv47991 Human pro
30	127.6	8.3	482	10	Abt40796 Toxicity
31	126.6	8.2	1026	6	Abq73061 Human kin
32	126.6	8.2	1026	9	Aal56806 DNA encod
33	126.6	8.2	1026	10	Adb66785 DNA encod
34	124.2	8.1	1011	6	Abq73062 Human kin
35	124.2	8.1	1011	9	Aal56807 DNA encod
36	124.2	8.1	1011	10	Adb66787 DNA encod
37	121.8	7.9	3661	6	Aad28568 Human kin
38	121.8	7.9	3694	10	Adc10189 Human NOV
39	121.4	7.9	1014	6	Abx94615 Human HsK
40	121.2	7.9	3570	12	Adq19012 Human sof
41	119.8	7.8	2994	4	Abi28591 Drosophil
42	119.4	7.8	2034	4	Abi16733 Drosophil
43	119.4	7.8	4034	4	Abi16732 Drosophil
44	119.2	7.8	2132	10	Adc30338 Human nov
45	117	7.6	1839	10	Adc30569 Human nov

ALIGNMENTS

RESULT 1
ADc23339
ID ADc23339 standard; DNA; 1538 BP.
XX
AC ADc23339;
XX
DT 18-DEC-2003 (first entry)
XX
DE DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 3).
XX
KW human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;
KW cytosolic; cardiant; immunomodulator; antiinflammatory; gene therapy;
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;
KW cardiac hypertrophy; immune disorder; inflammation.
XX
OS Homo sapiens.
XX
FH Key
CDS Location/Qualifiers
FT 1..1538
FT /*tag= a
FT /product= "KID protein"
FT /transl_except= [pos: 4..5; aa: Pro)
FT /note= "This codon has an apparent 1 nucleotide deletion
that alters the reading frame"

US6387644-B1.
14-MAY-2002.
28-NOV-2000; 2000US-00724224.
20-APR-1999; 99US-00295612.
20-JUN-2000; 2000US-00597292.
(CYTO-) CYTOKINETICS INC.
Beraud C;
WPI; 2003-706919/67.
P-PSDB; ADc23340.
Identifying a candidate agent as modulator of function of a target
protein for treating cellular proliferation disorders by adding a
candidate agent to a mixture of the target protein that

OS Homo sapiens.
 XX US6762043-B1.
 XX 13-JUL-2004.
 XX 06-MAR-2002; 2002US-00093317.
 XX 20-APR-1999; 99US-00295612.
 XX 20-JUN-2000; 2000US-00597292.
 XX 28-NOV-2000; 2000US-00724224.
 XX (CYTO-) CYTOKINETICS INC.
 XX Beraud C;
 XX WPI; 2004-532491/51.
 XX New isolated microtubule motor protein, useful for screening modulators
 XX for treating cellular proliferation disorders such as cancer,
 XX hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
 XX inflammation.
 XX Disclosure; SEQ ID NO 3; 26pp; English.
 XX The invention relates to human microtubule motor proteins and the nucleic
 XX acids encoding them. The invention also relates to a method of screening
 XX for modulators of a motor protein which has microtubule stimulated ATPase
 XX activity, a method of testing for ATPase activity of microtubule motor
 XX proteins, methods to identify candidate agents that bind to a target
 XX protein or act as a modulator of the binding characteristics or
 XX biological activity of a target protein, modulators of the target
 XX protein, and methods of treating cellular proliferation disorders such as
 XX cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
 XX and inflammation, for treating disorders associated with kinesin-like DNA
 XX binding protein (KID) and for inhibiting KID. The sequences are used for
 XX screening for modulators of motor proteins useful for treating cellular
 XX proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
 XX hypertrophy, immune disorders and inflammation, for treating disorders
 XX associated with KID and for inhibiting KID and for treating autoimmune
 XX diseases, arthritis, graft rejection, inflammatory bowel disease and
 XX proliferation induced after medical procedures including surgery and
 XX angioplasty. This sequence represents DNA encoding a human microtubule
 XX motor protein of the invention. Note: The specification states that this
 XX sequence encodes the protein featured as SEQ ID NO:4, but this does not
 XX appear to be the case.

Query Match
 Best Local Similarity 100.0%; Score 1538; DB 12; Length 1538;
 Matches 1538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCAGCGCGCGGCTCGACGACGAGCGAGCGCGAGATGCGGCGAGCTTCAGCGG 60
 DB 1 ATGCAGCGCGCGGCTCGACGACGAGCGAGCGCGAGATGCGGCGAGCTTCAGCGG 60
 QY 61 CGGCGATCTCAGGAGCTGCTGCTCGCTAAGCAAGATTGGAGCTACTCGTCTCCAC 120
 DB 61 CGGCGATCTCAGGAGCTGCTGCTCGCTAAGCAAGATTGGAGCTACTCGTCTCCAC 120
 QY 121 CTCACAGCTCGCTAAGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 DB 121 CTCACAGCTCGCTAAGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 QY 181 CAAGTGATCCCCCTGTGTGCGGGGATGGAAGCTGCTCTCTAGAGATTGCTAACTGGA 240
 DB 181 CAAGTGATCCCCCTGTGTGCGGGGATGGAAGCTGCTCTCTAGAGATTGCTAACTGGA 240
 QY 241 GGAACACAGAGAGACTCTCAAAATACAGTTGATGCTCTCTATGGGAGAGGAGTACTC 300
 DB 241 GGAACACAGAGAGACTCTCAAAATACAGTTGATGCTCTCTATGGGAGAGGAGTACTC 300

QY 301 AGCAGGACATCTATGTCAGGTTTCAGTGCAGCCATCTTAAGGACCTTGTCTGGAAGGGCAGA 360
 DB 301 AGCAGGACATCTATGTCAGGTTTCAGTGCAGCCATCTTAAGGACCTTGTCTGGAAGGGCAGA 360
 QY 361 ATGCCAGTGTGCTTGCCTATGACCCACAGGAGCTGGGAGAGCACAACAATCTCTGGGCA 420
 DB 361 ATGCCAGTGTGCTTGCCTATGACCCACAGGAGCTGGGAGAGCACAACAATCTCTGGGCA 420
 QY 421 GCCCAGAGCAACTCGGGGTGATCCCGGGGCTCTCATGACCTCTCTGCAAGGG 480
 DB 421 GCCCAGAGCAACTCGGGGTGATCCCGGGGCTCTCATGACCTCTCTGCAAGGG 480
 QY 481 AGGAGGGTCCCGAGGGCCCGGCTCTTGTGTACCATGTCTTACCTAGAGATCT 540
 DB 481 AGGAGGGTCCCGAGGGCCCGGCTCTTGTGTACCATGTCTTACCTAGAGATCT 540
 QY 541 ACCAGAGAGGATTTAGACCTCTTGGACCTCTGCTTGGGAGACCTTGGTAATCCGAGAAG 600
 DB 541 ACCAGAGAGGATTTAGACCTCTTGGACCTCTGCTTGGGAGACCTTGGTAATCCGAGAAG 600
 QY 601 ACTGCGCGGGGATATCTCTGATTCGGGTCTCTCCAGAGAGCCCATCAGTAGCTTTGCTG 660
 DB 601 ACTGCGCGGGGATATCTCTGATTCGGGTCTCTCCAGAGAGCCCATCAGTAGCTTTGCTG 660
 QY 661 ATTTTGTAGCGGCACTTCTCTGCGAGCCAGTCCGAAATCGGACTGTAGGAGCCACCCGCTCA 720
 DB 661 ATTTTGTAGCGGCACTTCTCTGCGAGCCAGTCCGAAATCGGACTGTAGGAGCCACCCGCTCA 720
 QY 721 ACCAGCGCTCTCCCGCAGTATGCTGCTCTCTGCTCAAGGTGGACCGGGAACGTT 780
 DB 721 ACCAGCGCTCTCCCGCAGTATGCTGCTCTCTGCTCAAGGTGGACCGGGAACGTT 780
 QY 781 TGGCCCCATTTCCGCGAGGAGGAGGAAACTCTACTGATGACTGTTGGTGGGTGAGAGG 840
 DB 781 TGGCCCCATTTCCGCGAGGAGGAGGAAACTCTACTGATGACTGTTGGTGGGTGAGAGG 840
 QY 841 ACNACCGCGCAGCAGGCAACAGGCGCTTGGCTTAAAGAGAGTGGAGCCATCAACACCT 900
 DB 841 ACNACCGCGCAGCAGGCAACAGGCGCTTGGCTTAAAGAGAGTGGAGCCATCAACACCT 900
 QY 901 CCTGTTTGTCTTGGGCAAGTGTGATGCGCTGAATCAGGCGCTCTCTGCTGTACCTT 960
 DB 901 CCTGTTTGTCTTGGGCAAGTGTGATGCGCTGAATCAGGCGCTCTCTGCTGTACCTT 960
 QY 961 ATCGGGACAGCAAGCTCACTGCTGCTTATGAGGACTCTCTGCTGGTGGCTCAGCCCACTA 1020
 DB 961 ATCGGGACAGCAAGCTCACTGCTGCTTATGAGGACTCTCTGCTGGTGGCTCAGCCCACTA 1020
 QY 1021 TCCTTTATTTGCCAACATTCGCCCTGAGAGACGCTTCTAGTACACAGTCTCCGCACTCA 1080
 DB 1021 TCCTTTATTTGCCAACATTCGCCCTGAGAGACGCTTCTAGTACACAGTCTCCGCACTCA 1080
 QY 1081 ACTTTGCTGCCAGGTCCCAAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTCGACG 1140
 DB 1081 ACTTTGCTGCCAGGTCCCAAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTCGACG 1140
 QY 1141 CTATGCTCTGGGACCTGTTAAGTGTCTCAAGAAATGCTTGGTCCACAGAGGCA 1200
 DB 1141 CTATGCTCTGGGACCTGTTAAGTGTCTCAAGAAATGCTTGGTCCACAGAGGCA 1200
 QY 1201 AGAGAGCCGAGGCGCTGAGGAGAGGAGATTGGGAGCCCTGAGCCCATGGCAGCTCCAG 1260
 DB 1201 AGAGAGCCGAGGCGCTGAGGAGAGGAGATTGGGAGCCCTGAGCCCATGGCAGCTCCAG 1260
 QY 1261 CCTTGTCTCCAGAACTCAGACCCCTTACAGAGCTAAGCAGCATGGAACCGGCGCATGC 1320
 DB 1261 CCTTGTCTCCAGAACTCAGACCCCTTACAGAGCTAAGCAGCATGGAACCGGCGCATGC 1320
 QY 1321 TGGAGCGCTCTCTCAGCTTTGGAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
 DB 1321 TGGAGCGCTCTCTCAGCTTTGGAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
 QY 1381 TGTGTAGTACCCCAAGCGAGCGGATGCTGCTTAATGAAGACAGTAGAAGAGGAGACC 1440

QY 606 CGGGGGAATATCTGATTCGGGCTCTCCCAAGAGCCCATCAGTAGCTTTGCTGATTTT 665
 Db 628 CGGGGGAATATCTGATTCGGGCTCTCCCAAGAGCCCATCAGTAGCTTTGCTGATTTT 687
 QY 666 GAGGGGCACTCTGCGAGCCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACAG 725
 Db 688 GAGGGGCACTCTCTGCCAGCCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACAG 747
 QY 726 CGCTCTCTCCGAGTCATGCTGTGCTCTCTGCTCAAGGTGGACCGGGAAGTTTGCC 785
 Db 748 CGCTCTCTCCGAGTCATGCTGTGCTCTCTGCTCAAGGTGGACCGGGAAGTTTGCC 807
 QY 786 CCAATTCGCGAGGAGAGGAAAATCTCACTGATTTGACTTGGCTGGGTGAGAGCAAC 845
 Db 808 CCAATTCGCGAGGAGAGGAAAATCTCACTGATTTGACTTGGCTGGGTGAGAGCAAC 867
 QY 846 CGGGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 905
 Db 868 CGGGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 927
 QY 906 TTTGCTCTGGGCAAGTGTAGATCGCTGAATCAGGGCTCCCTCGTGTACCTTATCGG 965
 Db 928 TTTGCTCTGGGCAAGTGTAGATCGCTGAATCAGGGCTCCCTCGTGTACCTTATCGG 987
 QY 966 GACAGCAAGCTCACTCGCTTATTCAGGACTCTCTGGGTGGGTGAGGAGGAGGAGGAG 1025
 Db 988 GACAGCAAGCTCACTCGCTTATTCAGGACTCTCTGGGTGGGTGAGGAGGAGGAGGAG 1047
 QY 1026 ATTGCGACATGTCCTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1085
 Db 1048 ATTGCGACATGTCCTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1107
 QY 1086 GCTGCGAGTCCAGAGGAGTGAATCAATCGGCTTTTACCAATGAGAGGCTCGAGCTCAT 1145
 Db 1108 GCTGCGAGTCCAGAGGAGTGAATCAATCGGCTTTTACCAATGAGAGGCTCGAGCTCAT 1167
 QY 1146 GCCTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTTGTTGTCACAGAGGCAAGAGA 1205
 Db 1168 GCCTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTTGTTGTCACAGAGGCAAGAGA 1227
 QY 1206 GCGGAGGCTCTGAGAGAGGAGATTTGGAGGCTCTGAGCCATGAGGAGCTTCCAGCTCT 1265
 Db 1228 GCGGAGGCTCTGAGAGAGGAGATTTGGAGGCTCTGAGCCATGAGGAGCTTCCAGCTCT 1287
 QY 1266 GCCTTCCAGAAATCAGGCTCTCAGAGGCTTAAGCAGCATGAGCCGCGCATCTGGAG 1325
 Db 1288 GCCTTCCAGAAATCAGGCTCTCAGAGGCTTAAGCAGCATGAGCCGCGCATCTGGAG 1347
 QY 1326 CGCTCTCTCAGCTTGGACCGTCTGCTTCCCTCCAGGGGAGCCAGGGGCGCTCTGTG 1385
 Db 1348 CGCTCTCTCAGCTTGGACCGTCTGCTTCCCTCCAGGGGAGCCAGGGGCGCTCTGTG 1407
 QY 1386 AGTACCCCAAGAGGAGGAGGAGTGTGCTTAATGAAGACAGTAGAGAGAGGAGCTTAGAG 1445
 Db 1408 AGTACCCCAAGAGGAGGAGGAGTGTGCTTAATGAAGACAGTAGAGAGAGGAGCTTAGAG 1467
 QY 1446 ATTGAGAGGCTTAAGAGCAAGCAAGAAAGAACTGGAGGCTCAAGATTTGGCCCAAGAGCT 1505
 Db 1468 ATTGAGAGGCTTAAGAGCAAGCAAGAAAGAACTGGAGGCTCAAGATTTGGCCCAAGAGCT 1527
 QY 1506 GAGGAAAGGAGAGCAATTTGTCACCAATG 1535
 Db 1528 GAGGAAAGGAGAGCAATTTGTCACCAATG 1557

RESULT 4

ID ADK40995

XX ADK40995 standard; DNA; 2099 BP.

AC ADK40995;

XX 06-MAY-2004 (first entry)

DT

XX

Novel human kinase gene #15.

cytostatic; immunomodulator; cardiant; neuroprotective; nootropic;
 aniparkinsonian; virucide; antibacterial; fungicide; ophthalmological;
 analgesic; hypotensive; immunosuppressive; kinase inhibitor; kinase;
 cancer; peripheral nervous system; central nervous system;
 Alzheimer's disease; Parkinson's disease; multiple sclerosis;
 amyotrophic lateral sclerosis; viral infection; prion infection;
 ocular disease; migraine; pain; sexual dysfunction; mood disorder;
 attention disorder; cognition disorder; hypotension; hypertension;
 psychotic disorder; neurological disorder; dyskinesia;
 metabolic disorder; organ transplant rejection; enzyme; gene; ds.

Homo sapiens.

W02003057841-A2.

17-JUL-2003.

31-DEC-2002; 2002WO-US041687.

31-DEC-2001; 2001US-0343169P.

(GRIG/) GRIGORIEV I V.
 (SUDA/) SUDARSANAM S.

Grigoriev IV, Sudarsanam S;

WPI; 2003-587115/55.

New isolated, enriched or purified nucleic acid molecule encoding a
 kinase polypeptide, useful for treating cancer, immune-related diseases,
 cardiovascular disease, brain or neuronal-associated diseases and
 metabolic disorders.

Claim 33; SEQ ID NO 102; 491pp; English.

The invention relates to novel isolated, enriched or purified nucleic acid
 molecules encoding a kinase polypeptide. The nucleic acid molecule
 comprises a sequence that: (a) encodes a kinase polypeptide; (b) is a
 complement of (a); (c) hybridizes under stringent conditions to (a) and
 encodes a naturally occurring kinase polypeptide; (d) encodes the
 polypeptide in (a), except that it lacks one or more, but not all, of an
 N-terminal domain, C-terminal catalytic domain, a catalytic region, a C-
 terminal domain, a coiled-coil structure region, a spacer region and a C-
 terminal tail; or (e) is a complement of (d). The nucleic acid molecules,
 polypeptides, methods and substance are useful for treating cancers,
 immune-related diseases or disorders, and metabolic disorders. The disorders are
 preferably cancers of the tissues or of hematopoietic origin, diseases of
 the central or peripheral nervous system, Alzheimer's disease, Parkinson's
 disease, multiple sclerosis, amyotrophic lateral sclerosis, viral
 infections, infections caused by prions, infections caused by
 bacteria, infections caused by fungi, ocular diseases, migraines, pain,
 sexual dysfunction, mood disorders, attention disorders, cognition
 disorders, hypotension, hypertension, psychotic disorders, neurological
 disorders, dyskinesias, metabolic disorders and organ transplant
 rejection. This sequence corresponds to the DNA encoding one of the
 kinase polypeptides of the invention.

Sequence 2099 BP; 484 A; 592 C; 628 G; 395 T; 0 U; 0 Other;

Query Match 98.8%; Score 1519; DB 10; Length 2099;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1530; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 6 GCCCGCGGCGGCTCGACGCGACGACGAGCGAGATGGCGGAGCTTCAGCGCGCGCG 65

Db 29 GCCCGCGGCGGCTCGACGCGACGACGAGCGAGATGGCGGAGCTTCAGCGCGCGCG 88

QY 66 ATCTCAGGAGCTGGTGGCTGTGGCTAAGCAAGATTGGAGCTACTCGTCGCTCCACCTCCA 125

Db 89 ATCTCAGGAGCTGGTGGCTGTGGCTAAGCAAGATTGGAGCTACTCGTCGCTCCACCTCCA 148

QY	126	GCTCGGTAAAGGTGGTGTGCGATCTGCGGCCATTTGTGATGGAAACACGCGGAGCAAGT	185
Db	149	GCTCGGTAAAGGTGGTGTGCGACTGCGGCCATTTGTGATGGAAACGCGGAGCAAGT	208
QY	186	GATCCCCCTGTGTGCGGGCATGGACAGTGTCTCTAGAGATTCTTAATCTGGAGAAC	245
Db	209	GATCCCCCTGTGTGCGGGCATGGACAGTGTCTCTAGAGATTCTTAATCTGGAGAAC	268
QY	246	CACAGAGAGACTCTCAAAATACCAAGTTTGTATGCCCTTCTATGGGAGAGGAGTACTCAGACG	305
Db	269	CACAGAGAGACTCTCAAAATACCAAGTTTGTATGCCCTTCTATGGGAGAGGAGTACTCAGACG	328
QY	306	GACATCTATGACGGTTCAGTGCAGCCCATCTTAAGGCACTTGTGTAAGGGGCAAGATGCC	365
Db	329	GACATCTATGACGGTTCAGTGCAGCCCATCTTAAGGCACTTGTGTAAGGGGCAAGATGCC	388
QY	366	AGTGTGCTTCCCTATGAGCCACAGAGCTGGGAAGCAGCACAAATGCTGGGCAGCCCA	425
Db	389	AGTGTGCTTCCCTATGAGCCACAGAGCTGGGAAGCAGCACAAATGCTGGGCAGCCCA	448
QY	426	GAGCAACCTGGGTGTATCCCGCGGGTCTCTATGGAACCTCTGAGCTCAAGGGAGGAG	485
Db	449	GAGCAACCTGGGTGTATCCCGCGGGTCTCTATGGAACCTCTGAGCTCAAGGGAGGAG	508
QY	486	GGTCCGAGGGCGGGCCATGGGCCCTTTCTGTCAACATGCTTACCTAGAGATCTACCG	545
Db	509	GGTCCGAGGGCGGGCCATGGGCCCTTTCTGTCAACATGCTTACCTAGAGATCTACCG	568
QY	546	GAGAGGTATTAGACCTCTGACCTGACCTGCTCGGAGACCTGTATATCGAGAAGACTGC	605
Db	569	GAGAGGTATTAGACCTCTGACCTGACCTGCTCGGAGACCTGTATATCGAGAAGACTGC	628
QY	606	CGGGGGAATATCTGTATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTT	665
Db	629	CGGGGGAATATCTGTATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTT	688
QY	666	GAGGGCACTTCTGCGCAGCAGTGCAGAAATCGACTGTAGAGCCACCGGCTCAACCAG	725
Db	689	GAGGGCACTTCTGCGCAGCAGTGCAGAAATCGACTGTAGAGCCACCGGCTCAACCAG	748
QY	726	CGCTCTCCGCGAGTCACTGTGCTCTGCTCAAGGTGGACGAGCGGGAACGTTTGGCC	785
Db	749	CGCTCTCCGCGAGTCACTGTGCTCTGCTCAAGGTGGACGAGCGGGAACGTTTGGCC	808
QY	786	CCATTTCCGCGAGCAGAGGGGAAACTCTA CTTGATTCAGTTCGCTGGGTGAGAGCAAC	845
Db	809	CCATTTCCGCGAGCAGAGGGGAAACTCTA CTTGATTCAGTTCGCTGGGTGAGAGCAAC	868
QY	846	CGGCGCACAGGCAACAAAGGGCTTTCGGCTAAAGAGAGTGGGCCATCAACCTCCCTG	905
Db	869	CGGCGCACAGGCAACAAAGGGCTTTCGGCTAAAGAGAGTGGGCCATCAACCTCCCTG	928
QY	906	TTTGTCTCGGCAAGTGTAGATGCGTCAATCAGGGCCCTCCCTCGTGTACCTTATCGG	965
Db	929	TTTGTCTCGGCAAGTGTAGATGCGTCAATCAGGGCCCTCCCTCGTGTACCTTATCGG	988
QY	966	GACAGCAAGCTCATCTGCGCTTATGGAAGTCTCTGCTGGGTGGCTCAGCCCAAGTATCCTT	1025
Db	989	GACAGCAAGCTCATCTGCGCTTATGGAAGTCTCTGCTGGGTGGCTCAGCCCAAGTATCCTT	1048
QY	1026	ATTGCCACATTTGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT	1085
Db	1049	ATTGCCACATTTGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT	1108
QY	1086	GCTGCGAGTCCAAAGGAGTGTATCAATCGGCCCTTTTACCATAGAGCCTGCAACCTCAT	1145
Db	1109	GCTGCGAGTCCAAAGGAGTGTATCAATCGGCCCTTTTACCATAGAGCCTGCAACCTCAT	1168
QY	1146	GCCTTGGACCTGTTAAGCTGTCTCAGAAAGAAATGCTTGTGTCACAGAGGCAAGAGA	1205
Db	1169	GCCTTGGACCTGTTAAGCTGTCTCAGAAAGAAATGCTTGTGTCACAGAGGCAAGAGA	1228

RESULT 5

ADC23343 standard; DNA; 1464 BP.

AC ADC23343;

18-DEC-2003 (first entry)

DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 7).

human; gene; ds; motor domain; kinesin-like DNA binding protein; KID; cytoskeletal; cardiant; immunomodulator; antiinflammatory; gene therapy; cancer; hyperplasia; restenosis; cellular proliferation disorder; cardiac hypertrophy; immune disorder; inflammation.

Homo sapiens.

Key Location/Qualifiers

CDS 1..1464

FT /*tag= a

FT /product= "KID protein"

US6387644-B1.

14-MAY-2002.

28-NOV-2000; 2000US-00724224.

20-APR-1999; 99US-00295612.

20-JUN-2000; 2000US-00597292.

(CYTO-) CYTOKINETICS INC.

Beraud C;

WPI; 2003-706919/67.

P-PSDB; ADC23344.

Identifying a candidate agent as modulator of function of a target protein for treating cellular proliferation disorders by adding a candidate agent to a mixture of the target protein that directly/indirectly produces ADP or phosphate.

Disclosure; SEQ ID NO 7; 26pp; English.

This invention relates to a novel method for high throughput screening systems used to identify compounds for the treatment of cellular proliferation disorders. Specifically, it refers to candidate agents that

are capable of modulating the activity of target proteins having motor domains, such that the target protein directly or indirectly produces ADP or phosphate. Furthermore, this activity can be determined using fluorescence or absorbance readouts. The present invention describes a method that identifies modulators of the target protein, which is a kinesin-like DNA binding protein (known as KID) as cyclostatic, cardants, immunomodulators and antiinflammatory. Accordingly, through gene therapy, they can be used for the treatment of cancer, hyperplasia, restenosis, cardiac hypertrophy, immune disorders and inflammation. This polynucleotide sequence is human KID DNA (SeqID 7) encoding a full length KID enzyme of the invention.

Sequence 1464 BP; 342 A; 409 C; 428 G; 285 T; 0 U; 0 Other;

Query Match	95.0%	Score 1461;	DB 10;	Length 1464;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1461;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	78	GGTGGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTGCTCACTCCAGCTCGCGTAAGG	137	
DB	4	GGTGGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTGCTCACTCCAGCTCGCGTAAGG	63	
QY	138	GTGGCTGTCGACCTGCGGCTTGTGATGGAACAGCGGGAGCAAGTGATCCCTCTGT	197	
DB	64	GTGGCTGTCGACCTGCGGCTTGTGATGGAACAGCGGGAGCAAGTGATCCCTCTGT	123	
QY	198	GTGGGGGATCGACAGCTGCTCTCTAGAGATTGCTAACTGGAGAACACACAGGAGACT	257	
DB	124	GTGGGGGATCGACAGCTGCTCTCTAGAGATTGCTAACTGGAGAACACACAGGAGACT	183	
QY	258	CTCAAAATACCAAGTTGATGCTTCTATGGGAGAGAGTATCTACAGGACATCTATGCA	317	
DB	184	CTCAAAATACCAAGTTGATGCTTCTATGGGAGAGAGTATCTACAGGACATCTATGCA	243	
QY	318	GGTTCAGTCAGCCATCTTAAGGCACTTGTGTAAGGCGCAATGCGAGTGTGCTGCC	377	
DB	244	GGTTCAGTCAGCCATCTTAAGGCACTTGTGTAAGGCGCAATGCGAGTGTGCTGCC	303	
QY	378	TATGGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGCCACAGACCACTGGG	437	
DB	304	TATGGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGCCACAGACCACTGGG	363	
QY	438	GTGATCCCGGGCTCTCATGAGCTCTCTGAGCTCACAAGGAGAGGGTGCAGAGGC	497	
DB	364	GTGATCCCGGGCTCTCATGAGCTCTCTGAGCTCACAAGGAGAGGGTGCAGAGGC	423	
QY	498	CGGCATCGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTATTA	557	
DB	424	CGGCATCGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTATTA	483	
QY	558	GACCTCTGGACCTGCTTCCGAGACCTGGTAATCCGAGAGAGCTGCCGGGGAAATATC	617	
DB	484	GACCTCTGGACCTGCTTCCGAGACCTGGTAATCCGAGAGAGCTGCCGGGGAAATATC	543	
QY	618	CTGATTCGGGTCTCTCCAGAACCCATCAGTACGTTTGTGATTTTGACGGCACTTC	677	
DB	544	CTGATTCGGGTCTCTCCAGAACCCATCAGTACGTTTGTGATTTTGACGGCACTTC	603	
QY	678	CTGCGACGAGTGGAAATCGACTGTAGGACCCACCGGCTCAACAGCGCTCTCCCGC	737	
DB	604	CTGCGACGAGTGGAAATCGACTGTAGGACCCACCGGCTCAACAGCGCTCTCCCGC	663	
QY	738	AGTCATGCTGTGCTCCTGGTCAAGTGGACACGCGGAAACGTTTGGCCCATTTTGGCAG	797	
DB	664	AGTCATGCTGTGCTCCTGGTCAAGTGGACACGCGGAAACGTTTGGCCCATTTTGGCAG	723	
QY	798	CGAGAGGAAATCTTACCTGATTTGACTTGGCTGAGGACCAACCGCGCACAGGC	857	
DB	724	CGAGAGGAAATCTTACCTGATTTGACTTGGCTGAGGACCAACCGCGCACAGGC	783	
QY	858	AACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTGTTTGTCTGGGC	917	
DB	784	AACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTGTTTGTCTGGGC	843	

QY	918	AAAGTGGTAGATGGCTGAATCAGGGCTCCTCGTGTACCTTATCGGACAGCAAGCTC	977	
DB	844	AAAGTGGTAGATGGCTGAATCAGGGCTCCTCGTGTACCTTATCGGACAGCAAGCTC	903	
QY	978	ACTCGCTATTGCGAGGACTCTCTGGTGGCTCAGCCACAGCATCTTATGCGCAACATT	1037	
DB	904	ACTCGCTATTGCGAGGACTCTCTGGTGGCTCAGCCACAGCATCTTATGCGCAACATT	963	
QY	1038	CCCTCTGAGACGCTTCTACTAGACAGCTCTCCGACCTCAACTTGTGCTCCAGGTC	1097	
DB	964	CCCTCTGAGACGCTTCTACTAGACAGCTCTCCGACCTCAACTTGTGCTCCAGGTC	1023	
QY	1098	AAGGAGTGTATCAATCGGCTTTTACCAATGAGAGCCTGACGCTCTATGCTTGGGACT	1157	
DB	1024	AAGGAGTGTATCAATCGGCTTTTACCAATGAGAGCCTGACGCTCTATGCTTGGGACT	1083	
QY	1158	GTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGAGCCGAGCCCT	1217	
DB	1084	GTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGAGCCGAGCCCT	1143	
QY	1218	GAGGAGGAGGAGATTGGGAGCCCTGAGCCCATGGCAGCTCCAGCTCTGCTCCAGAAA	1277	
DB	1144	GAGGAGGAGGAGATTGGGAGCCCTGAGCCCATGGCAGCTCCAGCTCTGCTCCAGAAA	1203	
QY	1278	CTCAGCCCTTACAGAAAGCTAAGCAGCATGGACCCGATGCTGGAGCGCTCTCTCAGC	1337	
DB	1204	CTCAGCCCTTACAGAAAGCTAAGCAGCATGGACCCGATGCTGGAGCGCTCTCTCAGC	1263	
QY	1338	TTGAGCCGTCTGCTTGGCTTCCAGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG	1397	
DB	1264	TTGAGCCGTCTGCTTGGCTTCCAGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG	1323	
QY	1398	CGAGCGGAGTGTCTAAGACAGTAAAGAGAGGACCTAGAGATTGAGAGGCTT	1457	
DB	1324	CGAGCGGAGTGTGTCTAAGACAGTAAAGAGAGGACCTAGAGATTGAGAGGCTT	1383	
QY	1458	AAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1517	
DB	1384	AAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1443	
QY	1518	AACCATTTGCCACAATGTGA	1538	
DB	1444	AACCATTTGCCACAATGTGA	1464	
RESULT 6				
ID	ADQ60233	standard; DNA; 1464 BP.		
AC	ADQ60233;			
XX	23-SEP-2004	(first entry)		
XX	Human microtubule motor protein DNA #4.			
XX	Human microtubule motor protein; gene; ds;			
XX	cellular proliferation disorder; cancer; hyperplasia; restenosis;			
XX	cardiac hypertrophy; immune disorder; inflammation;			
XX	kinesin-like DNA binding protein; KID; autoimmune disease; arthritis;			
XX	graft rejection; inflammatory bowel disease; angioplasty.			
OS	Homo sapiens.			
XX	Key	Location/Qualifiers		
FT	CDS	1..1464		
FT		/*tag= a		
FT		/product= "Human microtubule motor protein #4"		
XX	US6762043-B1.			
XX	13-JUL-2004.			
XX				

QY 1331 CCT-CAGTTGGACCGTCTGCTTGGCTCCAGGGAGCCAGGGGCCCCCTCTCTGAGTA 1389
 DB 1320 CTNACGTTGGACCGTCTGCTTGGCTCCAGGGAGCCAGGGGCCCCCTCTCTGAGTA 1379
 QY 1390 CCCAAGCGAGAGCGGATGCTCTAATGAAGACAGTAGAAGAGAGGACCTAGAGATTG 1449
 DB 1380 CCCAAGCGAGAGCGGATGCTCTAATGAAGACAGTAGAAGAGAGGACCTAGAGATTG 1439
 QY 1450 AGAGGCTTAAGACGAGCAAAAGAACTGGAGGCCAAGATGTTGGGCCAGAGGCTGAGG 1509
 DB 1440 ARAGGCTTAAGACGAGCAAAAGAACTGGAGGCCAAGATGTTGGGCCAGAGGCTGAGG 1499
 QY 1510 AAAGGAGACCAATTGTCACCAATG 1535
 DB 1500 AAAGGAGACCAATTGTCACCAATG 1525

RESULT 8
 ADC23337
 ID ADC23337 standard; DNA; 1115 BP.
 XX AC ADC23337;
 XX XX
 DT 18-DEC-2003 (first entry)
 XX DE DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 1).
 XX human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;
 KW cytosolic; cardiant; immunomodulator; antiinflammatory; gene therapy;
 KW cancer; hyperplasia; restenosis; cellular proliferation disorder;
 KW cardiac hypertrophy; immune disorder; inflammation.
 XX OS Homo sapiens.
 XX FH Key
 FT CDS
 FT 1. .1115
 FT /tag= a
 FT /product= "KID protein"
 FT /transl_except= (pos: 1. .5; aa: Met)
 FT /notes= "This codon has an apparent 2 nucleotide insertion
 FT that alters the reading frame"

US6387644-B1.
 PD 14-MAY-2002.
 XX 28-NOV-2000; 2000US-00724224.
 XX PF 20-APR-1999; 99US-00295612.
 XX PR 20-JUN-2000; 2000US-00597292.
 XX XX
 XX (CYTO-) CYTOKINETICS INC.
 XX Beraud C;
 XX WPI; 2003-706919/67.
 XX P-PSDB; ADC23338.
 XX
 XX Identifying a candidate agent as modulator of function of a target
 PT protein for treating cellular proliferation disorders by adding a
 PT candidate agent to a mixture of the target protein that
 PT directly/indirectly produces ADP or phosphate.
 XX
 XX Disclosure; SEQ ID NO 1; 26pp; English.

XX This invention relates to a novel method for high throughput screening
 CC systems used to identify compounds for the treatment of cellular
 CC proliferation disorders. Specifically, it refers to candidate agents that
 CC are capable of modulating the activity of target proteins having motor
 CC domains, such that the target protein directly or indirectly produces ADP
 CC or phosphate. Furthermore, this activity can be determined using
 CC fluorescence or absorbance readouts. The present invention describes a

CC method that identifies modulators of the target protein, which is a
 CC kinesin-like DNA binding protein (known as KID) as cytosolic, cardiant,
 CC immunomodulators and antiinflammatory. Accordingly, through gene
 CC therapy, they can be used for the treatment of cancer, hyperplasias,
 CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This
 CC polynucleotide sequence is human KID DNA (SeqID 1) encoding a full length
 CC KID enzyme of the invention.
 XX
 SQ Sequence 1115 BP; 241 A; 318 C; 333 G; 223 T; 0 U; 0 Other;
 Query Match 72.3%; Score 1112.4; DB 10; Length 1115;
 Best Local Similarity 99.9%; Pred. No. 6.5e-296;
 Matches 1113; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGCAGCCCGGGCGGCTCGACGACAGAGCGGCGAGATGGCGGCGAGCTTCAGCGG 60
 DB 1 ATGCAGCCCGGGCGGCTCGACGACAGAGCGGCGAGATGGCGGCGAGCTTCAGCGG 60
 QY 61 CGGCGATCTCAGGAGCTGGTCTGCTGCTGAAGCAAGATTGGAGCTACTCGTCTCCAC 120
 DB 61 CGGCGATCTCAGGAGCTGGTCTGCTGCTGAAGCAAGATTGGAGCTACTCGTCTCCAC 120
 QY 121 CTCAGCTCGCTAAGGGTGGTCTGCGACTGCGGCATTTGFGATGGAACAGCGGAG 180
 DB 121 CTCAGCTCGCTAAGGGTGGTCTGCGACTGCGGCATTTGFGATGGAACAGCGGAG 180
 QY 181 CAACTGATCCCCCTGTGTCGGGCGATGGACAGCTGCTCTCTAGAGATTGCTAACTGGA 240
 DB 181 CAACTGATCCCCCTGTGTCGGGCGATGGACAGCTGCTCTCTAGAGATTGCTAACTGGA 240
 QY 241 GGAACACACAGGAGACTCTCAATACAGTTTGTATGCTTCTATGGGAGAGAGTACTC 300
 DB 241 GGAACACACAGGAGACTCTCAATACAGTTTGTATGCTTCTATGGGAGAGAGTACTC 300
 QY 301 AGCAGGACATCTATGCAAGTTTCAGTGCAGCCCATCTTAAGGCACTTGTCTGGAAGGCGA 360
 DB 301 AGCAGGACATCTATGCAAGTTTCAGTGCAGCCCATCTTAAGGCACTTGTCTGGAAGGCGA 360
 QY 361 ATGCCAGTGTGCTTGCCTATGACCCACAGAGAGCTGGGAGAGCGCACATAGCTGGGCA 420
 DB 361 ATGCCAGTGTGCTTGCCTATGACCCACAGAGAGCTGGGAGAGCGCACATAGCTGGGCA 420
 QY 421 GCCAGAGCAACCTGGGGTGATCCCGGGGTCTCATGACCTCTCTGAGCTCACAAAGG 480
 DB 421 GCCAGAGCAACCTGGGGTGATCCCGGGGTCTCATGACCTCTCTGAGCTCACAAAGG 480
 QY 481 AGGAGGTGCCAGGGCCGGCCATGGGCCCTTCTGTCAACCATGTCTTACCTAGAGATCT 540
 DB 481 AGGAGGTGCCAGGGCCGGCCATGGGCCCTTCTGTCAACCATGTCTTACCTAGAGATCT 540
 QY 541 ACCAGGAGAGGTATTAGACCTCTGAGCCCTGCTTCCGGGAGACCTGTTAATCCGAGAG 600
 DB 541 ACCAGGAGAGGTATTAGACCTCTGAGCCCTGCTTCCGGGAGACCTGTTAATCCGAGAG 600
 QY 601 ACTGCGGGGGAATATCTCTGATTTCCGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTG 660
 DB 601 ACTGCGGGGGAATATCTCTGATTTCCGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTG 660
 QY 661 ATTTTGAAGCGGACATTTCTCTGCCAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCA 720
 DB 661 ATTTTGAAGCGGACATTTCTCTGCCAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCA 720
 QY 721 ACCAGGCTCTTCCCGGAGTCTGCTGCTCTCTGCTCAAGGTGGACCGAGCGGAGACGTT 780
 DB 721 ACCAGGCTCTTCCCGGAGTCTGCTGCTCTCTGCTCAAGGTGGACCGAGCGGAGACGTT 780
 QY 781 TGGCCCAATTTTCAGGAGCGAGGGGAAAATCTTACTCTGATTTGCTTGGCTGGGTCAAGG 840
 DB 781 TGGCCCAATTTTCAGGAGCGAGGGGAAAATCTTACTCTGATTTGCTTGGCTGGGTCAAGG 840
 QY 841 ACAACCGGCGCACAGCAACAAGGGGCTTGGCTTAAAGAGAGTGGAGGATCAACACCT 900
 DB 841 ACAACCGGCGCACAGCAACAAGGGGCTTGGCTTAAAGAGAGTGGAGGATCAACACCT 900

901 CCTGTTGCTCTGGGCAAGTGGTAGATGCGTGAATCAGGCGCTCCCTCGTGTACCTT 960
901 CCTGTTGCTCTGGGCAAGTGGTAGATGCGTGAATCAGGCGCTCCCTCGTGTACCTT 960
961 ATCGGACACAGCTCCTCGCTATTGCGAGACTCTCTGGGTGGCTCAGCCACAGTA 1020
961 ATCGGACACAGCTCCTCGCTATTGCGAGACTCTCTGGGTGGCTCAGCCACAGTA 1020
1021 TCCTTATTGCCAATTCCTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCA 1080
1021 TCCTTATTGCCAATTCCTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCA 1080
1081 ACTTTGCTGCAGGTCCTCAAGGAGTGATCAATCG 1114
1081 ACTTTGCTGCAGGTCCTCAAGGAGTGATCAATCG 1114

RESULT 10
ID ADC23341 standard; DNA; 1041 BP.
XX AC ADC23341;
XX DT 18-DEC-2003 (first entry)
XX DE DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 5).
XX KW human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;
XX KW cytosolic; cardiant; immunomodulator; antiinflammatory; gene therapy;
XX KW cancer; hyperplasia; restenosis; cellular proliferation disorder;
XX KW cardiac hypertrophy; immune disorder; inflammation.
XX OS Homo sapiens.

XX Key Location/Qualifiers
XX CDS 1..1041
XX FT /*tag= a
XX FT /product= "KID protein"
XX US6387644-B1.
XX PD 14-MAY-2002.
XX PF 28-NOV-2000; 2000US-00724224.
XX PR 20-APR-1999; 99US-00295612.
XX PR 20-JUN-2000; 2000US-00597292.
XX PA (CYTO-) CYTOKINETICS INC.
XX PI Beraud C;
XX WPI; 2003-706919/67.
XX DR P-PSDB; ADC23342.
XX XX Identifying a candidate agent as modulator of function of a target
XX PT protein for treating cellular proliferation disorders by adding a
XX PT candidate agent to a mixture of the target protein that
XX PT directly/indirectly produces ADP or phosphate.
XX PS Disclosure; SEQ ID NO 5; 26pp; English.

XX This invention relates to a novel method for high throughput screening
XX systems used to identify compounds for the treatment of cellular
XX proliferation disorders. Specifically, it refers to candidate agents that
XX are capable of modulating the activity of target proteins having motor
XX domains, such that the target protein directly or indirectly produces ADP
XX or phosphate. Furthermore, this activity can be determined using
XX fluorescence or absorbance readouts. The present invention describes a
XX method that identifies modulators of the target protein, which is a
XX kinesin-like DNA binding protein (known as KID) as cytosolic, cardiants,
XX immunomodulators and antiinflammatories. Accordingly, through gene

CC therapy, they can be used for the treatment of cancer, hyperplasias,
CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This
CC polynucleotide sequence is human KID DNA (SeqID 5) encoding a full length
CC KID enzyme of the invention.
XX
SQ Sequence 1041 BP; 228 A; 295 C; 302 G; 216 T; 0 U; 0 Other;
Query Match 67.3%; Score 1035.4; DB 10; Length 1041;
Best Local Similarity 99.9%; Pred. No. 1.1e-274;
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 78 GGTCTGCTGGCTAAGCAAGATTGGAGCTACTCTGCTGCACCTCCAGCTCGCTAAGG 137
DB 4 GGTCTGCTGGCTAAGCAAGATTGGAGCTACTCTGCTGCACCTCCAGCTCGCTAAGG 63
QY 138 GTGGTGTGCGACTGCGGCATTTTGGATGGAACAGCGGGAGCAAGTATCCCCCTGT 197
DB 64 GTGGTGTGCGACTGCGGCATTTTGGATGGAACAGCGGGAGCAAGTATCCCCCTGT 123
QY 198 GTGCGGGCATGACAGCTGCTCTCTAGAGATTGCTAACTGAGGAACACACAGGAGCT 257
DB 124 GTGCGGGCATGACAGCTGCTCTCTAGAGATTGCTAACTGAGGAACACACAGGAGCT 183
QY 258 CTCAAATACCACTTTGATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 317
DB 184 CTCAAATACCACTTTGATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 243
QY 318 GGTTCAGTCAGCCCATCTCTAAGGCACTTGTGGAGGGCAGAAATGCCAGTGTCTTGGC 377
DB 244 GGTTCAGTCAGCCCATCTCTAAGGCACTTGTGGAGGGCAGAAATGCCAGTGTCTTGGC 303
QY 378 TATGGACCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGGCCACAGCAACCTGGG 437
DB 304 TATGGACCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGGCCACAGCAACCTGGG 363
QY 438 GTGATCCCGGGCTCTCATGACCTCTCTGAGCTCAAGGAGGAGGCTGCCGAGGC 497
DB 364 GTGATCCCGGGCTCTCATGACCTCTCTGAGCTCAAGGAGGAGGCTGCCGAGGC 423
QY 498 CGGCCATGGGCTCTCTGACCTCTCTACCTAGAGATCTACAGGAGAGAGTATTTA 557
DB 424 CGGCCATGGGCTCTCTGACCTCTCTACCTAGAGATCTACAGGAGAGAGTATTTA 483
QY 558 GACCTCTGGACCTCTCTCGGAGACCTGTTAATCCGAGAGACTGCGGGGGAATATC 617
DB 484 GACCTCTGGACCTCTCTCGGAGACCTGTTAATCCGAGAGACTGCGGGGGAATATC 543
QY 618 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTTCAGCGGCACTTC 677
DB 544 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTTCAGCGGCACTTC 603
QY 678 CTGCCAGCCAGTCAAAATCGGAGCTGAGGAGCCACCGGCTCAACAGGCTCTCTCCCGC 737
DB 604 CTGCCAGCCAGTCAAAATCGGAGCTGAGGAGCCACCGGCTCAACAGGCTCTCTCCCGC 663
QY 738 AGTCATGCTGCTCTCTGCTCAAGTGGACACAGCGGAAACGTTTGGCCCATTTTCGCCAG 797
DB 664 AGTCATGCTGCTCTCTGCTCAAGTGGACACAGCGGAAACGTTTGGCCCATTTTCGCCAG 723
QY 798 CGAGAGGGAACCTCTACCTGATTTGGCTGGCTGAGGAGCAACCGCGCCACAGGC 857
DB 724 CGAGAGGGAACCTCTACCTGATTTGGCTGGCTGAGGAGCAACCGCGCCACAGGC 783
QY 858 AACAGGGCTCTCGCTTAAAGAGAGTGGAGCCATCAACACCTCTCTTGTGTCTGGGC 917
DB 784 AACAGGGCTCTCGCTTAAAGAGAGTGGAGCCATCAACACCTCTCTTGTGTCTGGGC 843
QY 918 AAAGTGTAGATCGCTGAATAGGGCTCTCCCTCGTGTACTTATCGGAGACAGCACTC 977
DB 844 AAAGTGTAGATCGCTGAATAGGGCTCTCCCTCGTGTACTTATCGGAGACAGCACTC 903
QY 978 ACTCGCTATTGAGGAGCTCTCTGGGTGGCTCAGGCCACAGTATCTTATTCACAACTT 1037

Db 904 ACTCGCTATTGACGACTCTCTGGTGGCTCAGCCGACAGATATCTTATTGCCAACATT 963
QY 1038 GCCCTCTGACAGCGCTTCTTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGGTCC 1097
Db 964 GCCCTCTGACAGCGCTTCTTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGGTCC 1023
QY 1098 AAGGAGGTGATCAATCG 1114
Db 1024 AAGGAGGTGATCAATG 1040

RESULT 11

ADQ60231
ID ADQ60231 standard; DNA; 1041 BP.

AC ADQ60231;

DT 23-SEP-2004 (first entry)

XX Human microtubule motor protein DNA #3.

XX Human; microtubule motor protein; gene; ds;
KW cellular proliferation disorder; cancer; hyperplasia; restenosis;
KW cardiac hypertrophy; immune disorder; inflammation;
KW kinesin-like DNA binding protein; KID; autoimmune disease; arthritis;
KW graft rejection; inflammatory bowel disease; angioplasty.

XX Homo sapiens.

XX Key Location/Qualifiers
FH CDS 1..1041

FT /*tag= a

FT /product= "Human microtubule motor protein #3"

XX US6762043-B1.

XX 13-JUL-2004.

XX 06-MAR-2002; 2002US-00093317.

XX 20-APR-1999; 99US-00295612.

XX 20-JUN-2000; 2000US-00597292.

XX 28-NOV-2000; 2000US-00724224.

XX (CYTO-) CYTOKINETICS INC.

XX Beraud C;

XX WPI; 2004-532491/51.

XX P-PSDB; ADQ60232.

XX New isolated microtubule motor protein, useful for screening modulators
PT for treating cellular proliferation disorders such as cancer,
PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
PT inflammation.

XX Disclosure; SEQ ID NO 5; 26pp; English.

XX The invention relates to human microtubule motor proteins and the nucleic
CC acids encoding them. The invention also relates to a method of screening
CC for modulators of a motor protein which has microtubule stimulated ATPase
CC activity, a method of testing for ATPase activity of microtubule motor
CC proteins, methods to identify candidate agents that bind to a target
CC protein or act as a modulator of the binding characteristics or
CC biological activity of a target protein, modulators of the target
CC protein, and methods of treating cellular proliferation disorders such as
CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders
CC and inflammation, for treating disorders associated with kinesin-like DNA
CC binding protein (KID) and for inhibiting KID. The sequences are used for
CC screening for modulators of motor proteins useful for treating cellular
CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac
CC hypertrophy, immune disorders and inflammation, for treating disorders
CC associated with KID and for inhibiting KID and for treating autoimmune

CC diseases, arthritis, graft rejection, inflammatory bowel disease and
CC proliferation induced after medical procedures including surgery and
CC angioplasty. This sequence represents DNA encoding a human microtubule
CC motor protein of the invention.

XX SQ Sequence 1041 BP; 228 A; 295 C; 302 G; 216 T; 0 U; 0 Other;

Query Match 67.3%; Score 1035.4; DB 12; Length 1041;
Best Local Similarity 99.9%; Pred. No. 1.le-274;
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 78 GGTGGCTGCGCTAAGCAAGATTGGAGCTACTCTGTCTCCACCTCCAGCTCCGTAAGG 137
Db 4 GGTGGCTGCGCTAAGCAAGATTGGAGCTACTCTGTCTCCACCTCCAGCTCCGTAAGG 63
QY 138 GTGGCTGTGCGACTGCGGCCATTTGTGATGGAACAGCGGGAGCAAGTGATCCCCCTGT 197
Db 64 GTGGCTGTGCGACTGCGGCCATTTGTGATGGAACAGCGGGAGCAAGTGATCCCCCTGT 123
QY 198 GTGGGGGCATGGACAGCTGCTCTCTAGAGATTGTCTAACTGGAGGAACACACAGGAGACT 257
Db 124 GTGGGGGCATGGACAGCTGCTCTCTAGAGATTGTCTAACTGGAGGAACACACAGGAGACT 183
QY 258 CTCAAAATACCAAGTTTGATGCCCTTCTATGGGAGAGAGTACTCAGCAGACATCTATGCA 317
Db 184 CTCAAAATACCAAGTTTGATGCCCTTCTATGGGAGAGAGTACTCAGCAGACATCTATGCA 243
QY 318 GGTTCAGTGCAGCCCATCTTAAGGCATTTGCTGGAAGGGCAGATGCCAGTGCTTGGC 377
Db 244 GGTTCAGTGCAGCCCATCTTAAGGCATTTGCTGGAAGGGCAGATGCCAGTGCTTGGC 303
QY 378 TATGACCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGCCAGAGCAACTGGG 437
Db 304 TATGACCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGCCAGAGCAACTGGG 363
QY 438 GTGATCCCGGGCTCTCATGGAACCTCTGTCAGCTCAAGAGGAGAGGGTGCAGAGGC 497
Db 364 GTGATCCCGGGCTCTCATGGAACCTCTGTCAGCTCAAGAGGAGAGGGTGCAGAGGC 423
QY 498 CGGCCATGGGCGCTTCTGTCAACCATGCTTACCTAGAGATCTACCAGAGAGGTATTA 557
Db 424 CGGCCATGGGCGCTTCTGTCAACCATGCTTACCTAGAGATCTACCAGAGAGGTATTA 483
QY 558 GACCTCTGGACCTGCTTGGGAGAGCTGTGTAATCCGAGAGAGCTCCGGGGGAATATC 617
Db 484 GACCTCTGGACCTGCTTGGGAGAGCTGTGTAATCCGAGAGAGCTCCGGGGGAATATC 543
QY 618 CTGATTCGGGTCTCTCCAGAACCCATCATAGTAGCTTTGTGATTTTGAAGCGGCACTTC 677
Db 544 CTGATTCGGGTCTCTCCAGAACCCATCATAGTAGCTTTGTGATTTTGAAGCGGCACTTC 603
QY 678 CTGCGAGCCAGTCGAATCCGAGCTGTAGGAGCCACCGGCTCAACAGCGCTCTCCCGC 737
Db 604 CTGCGAGCCAGTCGAATCCGAGCTGTAGGAGCCACCGGCTCAACAGCGCTCTCCCGC 663
QY 738 AGTCATGCTGTGCTCTGGTCAAGGTGGACCGGGAACCTTTGGGCCCATTTTGCACAG 797
Db 664 AGTCATGCTGTGCTCTGGTCAAGGTGGACCGGGAACCTTTGGGCCCATTTTGCACAG 723
QY 798 CGAGAGGAAATCTCTACCTGATTTGCTGGTTCAGAGGACAAACCGGCGCACAGGC 857
Db 724 CGAGAGGAAATCTCTACCTGATTTGCTGGTTCAGAGGACAAACCGGCGCACAGGC 783
QY 858 AACAAAGGCTCTCGCTAAAGAGAGTGGAGCCATCAACACTCCCTGTTTGTCTGGGC 917
Db 784 AACAAAGGCTCTCGCTAAAGAGAGTGGAGCCATCAACACTCCCTGTTTGTCTGGGC 843
QY 918 AAAGTGTAGATGGCTGAAATCAGGGCTCCCTCGTGTACTTATCGGGACAGCAAGCTC 977
Db 844 AAAGTGTAGATGGCTGAAATCAGGGCTCCCTCGTGTACTTATCGGGACAGCAAGCTC 903
QY 978 ACTCGCTATTGACGAGTCTCTCGGTGGTTCAGGCCCAAGTATCTTATTGCCAACATT 1037

Db 904 ACTCGCCTATTGAGGACTCTCTGGTGGTTCAGCCACACAGTATCCTTTATTGCCAACATT 963
QY 1038 GCCCCTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACATCACTTTGCTGCGAGTCC 1097
Db 964 GCCCCTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACATCACTTTGCTGCGAGTCC 1023
QY 1098 AAGGAGGTGATCAATCG 1114
Db 1024 AAGGAGGTGATCAATTG 1040

RESULT 12
ABK70260
ID ABK70260 standard; cDNA; 491 BP.
XX AC ABK70260;
DT 15-JUL-2002 (first entry)
XX Human lung cancer associated cDNA SEQ ID 131.
XX DE Human; ss; gene; lung cancer; cytostatic; tumour; vaccine.
XX KW Homo sapiens.
XX OS
XX PN WO200224057-A2.
XX PD 28-MAR-2002.
XX PF 20-SEP-2001; 2001WO-US042232.
XX PR 22-SEP-2000; 2000US-0234837P.
XX PR 10-OCT-2000; 2000US-0239440P.
XX PR 29-JUN-2001; 2001US-0301928P.
XX PA (CORI-) CORIXA CORP.
XX PI Benson DR, Mohamath R, Lodes MJ;
XX WPI; 2002-372001/40.
XX
XX New tumor lung proteins and nucleic acids encoding the proteins, useful
PT as vaccines and for treating, preventing, diagnosing or monitoring lung
PT cancer.
XX
XX Claim 1; Page 144; 189pp; English.

The invention relates to an isolated polynucleotide comprising a sequence
CC selected from 183 human DNA sequences (appearing as ABK70130-ABK70312),
CC or their fragments, homologues, variants or complements and their encoded
CC polypeptides. Also included are an expression vector comprising the
CC polynucleotide operably linked to an expression control sequence; a host
CC cell transformed or transfected with an expression vector of; an isolated
CC antibody, or its antigen-binding fragment that specifically binds to the
CC polypeptide; a method for detecting the presence of a cancer in a patient
CC; a fusion protein comprising at least the polypeptide; an
CC oligonucleotide that hybridizes to the polynucleotide under moderately
CC stringent conditions; a method for stimulating and/or expanding T cells
CC specific for a tumour protein; an isolated T cell population comprising T
CC cells prepared from the method of above; a composition comprising a first
CC component consisting of carriers and immunostimulants, and a second
CC component selected from the polynucleotides, proteins, antibodies, fusion
CC proteins, T cell populations and antigen presenting cells expressing the
CC polypeptide; methods for stimulating an immune response or treating
CC cancer in a patient by administering the composition and diagnostic kits
CC comprising at least one of the oligonucleotide of, or an antibody and a
CC detection reagent consisting of a reporter group. The polypeptides and
CC polynucleotides are useful as vaccines for the treatment or prevention of
CC lung cancer, and for diagnosis and monitoring of such cancer. The
CC polynucleotide, polypeptide and antigen presenting cells can be used to
CC stimulate or expand T cells specific for a tumorous protein. The
CC polynucleotides may be used as probes or primers for nucleic acid
CC hybridisation, and in the preparation of ribozyme molecules for

CC inhibiting expression of tumour polypeptides and proteins in tumour
CC cells. The present sequence is one of the 183 lung cancer associated
CC polynucleotides
XX
SQ Sequence 491 BP; 127 A; 140 C; 136 G; 88 T; 0 U; 0 Other;
Query Match 31.2%; Score 480; DB 6; Length 491;
Best Local Similarity 99.8%; Pred. No. 9.3e-122;
Matches 491; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1010 AGCCACAGTATCCTTTATTGCCAACATTTGCCCTGAGAGACGCTTCTTACCTAGACACAGT 1069
Db 1 AGCCACAGTATCCTTTATTGCCAACATTTGCCCTGAGAGACGCTTCTTACCTAGACACAGT 60
QY 1070 CTCGCGACTCAACTTTGCTGCCAGGTCCAGGAGGTGATCAATCGGCTTTTACCAATGA 1129
Db 61 CTCGCGACTCAACTTTGCTGCCAGGTCCAGGAGGTGATCAATCGGCTTTT-CCAATGA 119
QY 1130 GAGCCTGCGAGCCTCATGCTTGGGACCTGTTAAGCTGTCTCAGAAAAGAAATTTGTTGGTCC 1189
Db 120 GAGCCTGCGAGCCTCATGCTTGGGACCTGTTAAGCTGTCTCAGAAAAGAAATTTGTTGGTCC 179
QY 1190 ACCAGAGGCAAGAGAGAGCCCGAGGCGCTTCAGAGAGAGGAGATTGGGAGCCCTGAGGCCAT 1249
Db 180 ACCAGAGGCAAGAGAGAGCCCGAGGCGCTTCAGAGAGAGGAGATTGGGAGCCCTGAGGCCAT 239
QY 1250 GGCAGCTCAGCCTCTGCTCCAGAAAATCAGCCCCCTTACAGAAAGTAAAGAGCATGGA 1309
Db 240 GGCAGCTCAGCCTCTGCTCCAGAAAATCAGCCCCCTTACAGAAAGTAAAGAGCATGGA 299
QY 1310 CCGGGCCATGCTGGAGCGCTCTCAGCTTGGACCGCTGCTTGGCTTCCAGGGAGGCCA 1369
Db 300 CCGGGCCATGCTGGAGCGCTCTCAGCTTGGACCGCTGCTTGGCTTCCAGGGAGGCCA 359
QY 1370 GGGGGCCCTCTGTTGAGTACCCAAAGCGAGAGCGGATGGTCTTAATGAAGACAGTAGA 1429
Db 360 GGGGGCCCTCTGTTGAGTACCCAAAGCGAGAGCGGATGGTCTTAATGAAGACAGTAGA 419
QY 1430 AGAAGAGGACCTAGAGATTGAGAGGCTTAAGACGAAGCAAAAAGAACTGGAGGCCAAGAT 1489
Db 420 AGAAGAGGACCTAGAGATTGAGAGGCTTAAGACGAAGCAAAAAGAACTGGAGGCCAAGAT 479
QY 1490 GTTGGCCAGAA 1501
Db 480 GTTGGCCAGAA 491

RESULT 13
AAS38879
ID AAS38879 standard; cDNA; 386 BP.
XX AC AAS38879;
DT 17-DEC-2001 (first entry)
XX DE Novel human diagnostic and therapeutic gene #1937.
XX KW Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.
XX OS Homo sapiens.
XX PN WO200166753-A2.
XX PD 13-SEP-2001.
XX PF 09-MAR-2001; 2001WO-US007787.
XX PR 09-MAR-2000; 2000US-0188609P.
XX PA (CHIR) CHIRON CORP.
XX PA (HYSE-) HYSEQ INC.
XX PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;

PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Kassam A, Lamson G;
PI Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
XX Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;
XX WPI; 2001-530177/58.
XX
XX New polynucleotides and polypeptides, useful for diagnosis and treatment
PT of breast, lung and colon cancer.
XX
XX Claim 1; Page 1080; 1193pp; English.
XX
XX The invention relates to new polynucleotides and polypeptides, useful for
CC diagnosis and treatment of breast, lung and colon cancer. The sequences
CC can be used in detecting differentially expressed genes correlated with a
CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample derived from a
CC cell suspected of being cancerous. They can also be used to inhibit
CC tumour growth by modulating expression of a gene product. AAS36943-
CC AAS39338 represent novel human diagnostic and therapeutic coding
CC sequences of the invention
XX
XX Sequence 386 BP; 105 A; 102 C; 118 G; 61 T; 0 U; 0 Other;
XX
XX Query Match 24.5%; Score 376.4; DB 4; Length 386;
XX Best Local Similarity 98.4%; Pred. No. 3e-93;
XX Matches 380; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
XX 1147 CTTGGGACCTTAAAGCTGCTCAGAAAGATTCTGGTCCACAGAGCAAGAGAG 1206
XX 1 CGTTCTGTGGTTAAAGCTGCTCAGAAAGATTCTGGTCCACAGAGCAAGAGAG 60
XX
XX 1207 CCCGAGGCCCTGAGGAGAGAGATTGGGAGCCCTGAGCCCATGCGCCTCTG 1266
XX 61 CCCGAGGCCCTGAGGAGAGAGATTGGGAGCCCTGAGCCCATGCGCCTCTG 120
XX
XX 1267 CTTCCAGAACTCAGCCCTTACAGAGCTTAAGCAGCATGACCCGCCCATGTGGAGC 1326
XX 121 CTTCCAGAACTCAGCCCTTACAGAGCTTAAGCAGCATGACCCGCCCATGTGGAGC 180
XX
XX 1327 GCCTCCTCAGCTTGAAGCTGCTGCTTCCAGGGAGCCAGGGGCCCTCTGTGA 1386
XX 191 GCCTCCTCAGCTTGAAGCTGCTGCTTCCAGGGAGCCAGGGGCCCTCTGTGA 240
XX
XX 1387 GTACCCCAAGCGAGCGGATGTGTGTAATGAACAGACAGTGAAGAGACCTAGAGA 1446
XX 241 GTACCCCAAGCGAGCGGATGTGTGTAATGAACAGACAGTGAAGAGACCTAGAGA 300
XX
XX 1447 TTGAGAGCTTAAAGCAGAGCAAGAAAGAACTGGAGGCCAAGATGTTGCCAGAGGCTG 1506
XX 301 TTGAGAGCTTAAAGCAGAGCAAGAAAGAACTGGAGGCCAAGATGTTGCCAGAGGCTG 360
XX
XX 1507 AGGAAAGGAGACCATTTGTCACCA 1532
XX 361 AGGAAAGGAGACCATTTGTCACCA 386
XX
XX RESULT 14
XX ACH4493
XX ID ACH4493 standard; cDNA; 464 BP.
XX AC ACH4493;
XX AC
XX 13-OCT-2003 (first entry)
XX
XX Human foetal brain cDNA #5218.
XX
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX genome mapping; biodiversity; genetic disorder.
XX Homo sapiens.
XX
XX US2003073623-A1.
XX

PD 17-APR-2003.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX (DRMA/) DRMANAC R T.
XX (LABA/) LABAT I.
XX (STAC/) STACHE-CRAIN B.
XX (DICK/) DICKSON M C.
XX (JONE/) JONES L W.
XX
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 31705; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030073623
XX
XX Sequence 464 BP; 96 A; 136 C; 131 G; 96 T; 0 U; 5 Other;
XX
XX Query Match 17.2%; Score 265; DB 9; Length 464;
XX Best Local Similarity 77.6%; Pred. No. 1.7e-62;
XX Matches 333; Conservative 0; Mismatches 95; Indels 1; Gaps 1;
XX
XX 200 GCGGGGATGACAGCTGCTCTTAGAGATTGCTAACTGGAGGAACCAACAGGAGACTCT 259
XX 37 GCCTGGAATTCATATCGCGTTCGTAGAGTGTAAACGCGCCGAAACCAACCGATATCT 96
XX
XX 260 CAAATACAGTTTGTATGCTTCTATGGGGAGAGAGTACTCAGCAGGACATCTATGACAG 319
XX 97 CCCTTACCTGTGCTTTCCTTTTGACAGCCATAGAGGACTGCGCTGACATGTACTCTCC 156
XX
XX 320 TTCAGTCGAGCCCATCTTAAGGCATCTTCTGGAAGGGCAGATGCCAGTGTCTTGCCTA 379
XX 157 TTGACCGTGGACCGTGGCCACTGCCCACTCTTTCGTAAATAAGCGGGAGCTGGCCTA 216
XX
XX 380 TGGACCCACAGAGCTGGGAAGAGCGCACCAATCTGCGGACGCCAGAGCAACCTGGGGT 439
XX 217 CGGTGCTCTGAGAGCTGGCATTAC-ATCTCGATCTTGGCAGCCAGAGCAACCTGGGGT 275
XX
XX 440 GATCCCGGGGCTCTCATGACCTCTGACCTCACAAGGGAGAGAGGTGCGGAGGGCG 499
XX 276 GATCCCGGGGCTCTCATGACCTCTGACCTCACAAGGGAGAGAGGTGCGGAGGGCG 335
XX
XX 500 GCCATGGCCCTTCTGTACCATGTCTTACCTAGAGATCTACCAGGAGAGGTATTAGA 559
XX 336 GCCATGGCCCTTCTGTACCATGTCTTACCTAGAGATCTACCAGGAGAGGTATTAGA 395
XX
XX 560 CCTCCTGACCCCTCTTCCGGGAGACCTGGTAAATCCGAGAAAGACTGCCCGGGGAATATCT 619
XX

Db 396 CCTCTGGACCTTGCTGGGAGACCTTGGATCCGAGAGACTGCGGGGGAATATCCT 455

Qy 620 GATTCCGGG 628

Db 456 GATTCCGGG 464

RESULT 15

ACH77408

ID ACH77408 standard; DNA; 531 BP.

XX AC ACH77408;

XX 29-JUL-2004 (first entry)

XX Human genome derived single exon probe #10603.

DE Human; probe; es; gene expression; single exon probe; microarray;

XX alternative splicing event; genomic alteration.

XX Homo sapiens.

XX US2003194704-A1.

XX 16-OCT-2003.

XX 03-APR-2002; 2002US-00029386.

XX 03-APR-2002; 2002US-00029386.

XX (PENN/) PENN S G.

XX (RANK/) RANK D R.

XX (HANZ/) HANZEL D K.

XX Penn SG, Rank DR, Hanzel DK;

XX WPI; 2004-119264/12.

XX New human genome-derived single exon nucleic acid probes useful for human

PT gene expression analysis, for identifying or characterizing alternative

PT splicing events, for assessing genomic alterations or as tools for

PT surveying tissues.

XX Claim 15; SEQ ID NO 10603; 80pp; English.

XX The invention relates to a nucleic acid probe for measuring human gene

CC expression, comprising any of the 27,400 fully defined nucleotide

CC sequences in the specification, or their complements or fragments, and

CC encoding at least 8 amino acids of any of the 688 amino acid sequences

CC fully defined in the specification. The probe is a single exon probe that

CC hybridises under high stringency conditions to a nucleic acid molecule

CC expressed in human cells or tissues. Also included are a spatially-

CC addressable set of single exon nucleic acid probes for measuring human

CC gene expression (comprising a plurality of single exon nucleic acid

CC probes cited above, where each of the plurality of probes is separately

CC and addressably isolatable or amplifiable from the plurality), a single

CC exon microarray for measuring human gene expression, a method of

CC measuring human gene expression, a vector comprising the single exon

CC probe cited above, an ORF-encoded peptide comprising at least 8

CC contiguous amino acids of any of the above-mentioned amino acid

CC sequences (optionally with conservative amino acid substitutions), an

CC isolated antibody that binds specifically to a peptide cited above,

CC methods of selling and/or licensing single exon probes or microarrays to

CC a customer desiring to measure gene expression, a method of providing

CC human gene expression data by subscription, and a computer-readable

CC storage medium which contains a database having a plurality of records

CC (each record including data on the expression of a single exon probe

CC cited above. The probe, methods and apparatus are useful in gene

CC expression analysis. The probes may be used as tools for surveying

CC tissues to detect the presence of expressed messages that contain their

CC specific exon, or in constructing genome-derived single exon microarrays.

CC In addition, the probes are used in identifying and characterising

CC alternative splicing events, in detecting and characterising gross

CC alterations in the genomic locus that includes their exon, in assessing

CC smaller genomic alterations, in priming the synthesis of nucleic acids,

CC or in expressing the ORF-encoded peptide. The present sequence is a human

CC single exon probe of the invention. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from USPTO at

CC seqdata.uspto.gov/sequence.html?DocID=20030194704

XX

SQ Sequence 531 BP; 117 A; 153 C; 149 G; 112 T; 0 U; 0 Other;

Query Match 15.2%; Score 233.2; DB 12; Length 531;

Best Local Similarity 98.7%; Pred. No. 1e-53;

Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 758 CAAGGTGGACCGGGAACGTTTGGCCCCCATTTCCGACGAGAGGAAAACCTTACCT 817

Db 34 CCAGGTGGACCGGGAACGTTTGGCCCCCATTTCCGACGAGAGGAAAACCTTACCT 93

Qy 818 GATTGACTTGGCTGGGTTCAGAGGACAAACCGGGCCACAGGCAACAAGGGCTTCGGCTAAA 877

Db 94 GATTGACTTGGCTGGGTTCAGAGGACAAACCGGGCCACAGGCAACAAGGGCTTCGGCTAAA 153

Qy 878 AGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAAAGTGTAGATCGCTGAA 937

Db 154 AGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAAAGTGTAGATCGCTGAA 213

Qy 938 TCAGGGCCCTCCCTCGTGTACCTTATCGGACAGCAAGCTCCTCGCTTATTCGAGGAC 995

Db 214 TCAGGGCCCTCCCTCGTGTACCTTATCGGACAGCAAGCTCCTCGCTTATTCGAGGTC 271

Search completed: November 10, 2004, 06:57:53

Job time : 798.24 secs

[illegible]

```

RESULT 2
US-10-093-317-3
; Sequence 3, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093.317
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: Human
US-10-093-317-3

```

Query Match	100.0%;	Score 1538;	DB 4;	Length 1538;
Best Local Similarity	100.0%;	Prod. No. 0;		
Matches 1538;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGCAGCCGGGGCGCTCGACGACGACAGAGGCGACGAGATCGCGCGAGCTTCAGCGG	60	
DB	1	ATGCAGCCGGGGCGCTCGACGACGACAGAGGCGACGCGAGATGCGCGAGCTTCAGCGG	60	
QY	61	CGGCGATCTCAGGAGCTGCTCGCTGCTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCAC	120	
DB	61	CGGCGATCTCAGGAGCTGCTCGCTGCTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCAC	120	
QY	121	CTTCAGCTCCGTAAGGGTGGCTGTGCGACTGCGGGCCATTGTGGATGGAAACACGGGGAG	180	
DB	121	CTTCAGCTCCGTAAGGGTGGCTGTGCGACTGCGGGCCATTGTGGATGGAAACACGGGGAG	180	
QY	181	CAAGTGATCCCCCTGTGTGCGGGGCATGGACAGCTGCTCTTAGAGATTGCTAACTGGA	240	
DB	181	CAAGTGATCCCCCTGTGTGCGGGGCATGGACAGCTGCTCTTAGAGATTGCTAACTGGA	240	
QY	241	GGAAACCAACGAGAGACTCTCAATACAGTTTGNATGCTTCTATGCGGAGAGGATGACTC	300	
DB	241	GGAAACCAACGAGAGACTCTCAATACAGTTTGNATGCTTCTATGCGGAGAGGATGACTC	300	
QY	301	AGCAGGACATCTATGCAGGTTTCAGTGACAGCCCATCTTAAGGCATCTGCTGGAAGGGCAGA	360	
DB	301	AGCAGGACATCTATGCAGGTTTCAGTGACAGCCCATCTTAAGGCATCTGCTGGAAGGGCAGA	360	
QY	361	ATGCCAGTGCTTGTCCTATGAGACCCACAGAGAGCTGGGAAAGACGACACAATGCTGGGCA	420	
DB	361	ATGCCAGTGCTTGTCCTATGAGACCCACAGAGAGCTGGGAAAGACGACACAATGCTGGGCA	420	
QY	421	GCCCAGAGCAACCTGGGGTGATCCCGGGGCTCTCATGGACCTCTCGAGCTCACAAAGG	480	
DB	421	GCCCAGAGCAACCTGGGGTGATCCCGGGGCTCTCATGGACCTCTCGAGCTCACAAAGG	480	
QY	481	AGGAGGTGCGAGGGCGGCGCATATGGGCCCTTTCTGTCACATGCTTTACTAGAGATCT	540	
DB	481	AGGAGGTGCGAGGGCGGCGCATATGGGCCCTTTCTGTCACATGCTTTACTAGAGATCT	540	
QY	541	ACCAGGAGAGGTAATTAGACTCTCTGACCTGCTTCGGGAGACCTGGTAATCCGAGAG	600	
DB	541	ACCAGGAGAGGTAATTAGACTCTCTGACCTGCTTCGGGAGACCTGGTAATCCGAGAG	600	
QY	601	ACTGCCGGGGGAATATCTCTGATTCGGGGTCTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTG	660	

786 CCATTTCGCCAGCAGAGGAGGAAACTCTACCTGATTGCTTGGCTGGCTCAGAGGCAAC 845
 Db CCAATTTGCCAGCAGAGGAGGAAACTCTACCTGATTGCTTGGCTGGCTCAGAGGCAAC 867
 Qy CGGCGCAGCAGCAACAAGGGCTTTTGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTG 905
 Db CGGCGCAGCAGCAACAAGGGCTTTTGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTG 927
 Qy TTGTCTCTGGCCAAAGTGTAGATGCGTGATCAGAGGCTTCCCTGCTGTACCTTATCGG 965
 Db TTGTCTCTGGCCAAAGTGTAGATGCGTGATCAGAGGCTTCCCTGCTGTACCTTATCGG 987
 Qy GACAGCAAGCTCACTCGCTTATTCAGAGCTCTCTGGTGGCTCAGGCCACAGTATCCTT 1025
 Db GACAGCAAGCTCACTCGCTTATTCAGAGCTCTCTGGTGGCTCAGGCCACAGTATCCTT 1047
 Qy ATTGCCAACATTGCCCCCTGAGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT 1085
 Db ATTGCCAACATTGCCCCCTGAGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT 1107
 Qy GCTGCCAGGTTCCAGAGAGTGTATCAATCGGCTTTTACCAATGAGAGCTTCCGACCTCAT 1145
 Db GCTGCCAGGTTCCAGAGAGTGTATCAATCGGCTTTTACCAATGAGAGCTTCCGACCTCAT 1167
 Qy GCCTTGGGACTGTTAAGCTGTCTCAGAAAGATTGCTTGGTCCACAGAGGCAAGAGA 1205
 Db GCCTTGGGACTGTTAAGCTGTCTCAGAAAGATTGCTTGGTCCACAGAGGCAAGAGA 1227
 Qy GCCGAGGCTCTGAGGAAGGAGATTGGAGGCTTGGAGCCCTAGGAGCTTCCAGGCTCT 1265
 Db GCCGAGGCTCTGAGGAAGGAGATTGGAGGCTTGGAGCCCTAGGAGCTTCCAGGCTCT 1287
 Qy GCCTCCAGAACTCAGCCCTTACAGAGCTAAGAGAGTGCATCCGCGCATCTGGAG 1325
 Db GCCTCCAGAACTCAGCCCTTACAGAGCTAAGAGAGTGCATCCGCGCATCTGGAG 1347
 Qy CGCTCTCAGCTTGGACCGCTCTGCTTCCCTCCAGGGAGCCAGGGGGCCCTCTGTTG 1385
 Db CGCTCTCAGCTTGGACCGCTCTGCTTCCCTCCAGGGAGCCAGGGGGCCCTCTGTTG 1407
 Qy AGTACCCCAAAGCAGAGCGGATGCTTAATGAAGACAGTGAAGAGAGGACCTAGAG 1445
 Db AGTACCCCAAAGCAGAGCGGATGCTTAATGAAGACAGTGAAGAGAGGACCTAGAG 1467
 Qy ATTGAGAGGCTTAAGACGAGCAAAAGAACTGGAGCCCAAGATGTTGGCCCAAGAGGCT 1505
 Db ATTGAGAGGCTTAAGACGAGCAAAAGAACTGGAGCCCAAGATGTTGGCCCAAGAGGCT 1527
 Qy GAGGAAAGGAGAACCAATTGTCCCAATG 1535
 Db GAGGAAAGGAGAACCAATTGTCCCAATG 1557

RESULT 4

US-09-724-224-7
 ; Sequence 7, Application US/09724224
 ; Patent No. 6387644
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
 ; FILE OF INVENTION: their use
 ; FILE REFERENCE: 1044
 ; CURRENT APPLICATION NUMBER: US/09/724,224
 ; PRIOR FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: 09/597,292
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 7
 ; LENGTH: 1464
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-724-224-7

Query Match 95.0%; Score 1461; DB 3; Length 1464;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 78 GGTGCGCTGTCGGCTAAAGCAAGATTGGAGCTACTCTGCTCCACCTCCAGCTCGCGTAAGG 137
 Db 4 GGTGCGCTGTCGGCTAAAGCAAGATTGGAGCTACTCTGCTCCACCTCCAGCTCGCGTAAGG 63
 Qy 138 GTGGCTGTGCGACTGCGGCCAATTGTGGATGGAACAGCGGAGCAAGTGTATCCCCCTGT 197
 Db 64 GTGGCTGTGCGACTGCGGCCAATTGTGGATGGAACAGCGGAGCAAGTGTATCCCCCTGT 123
 Qy 198 GTGCGGGGCATGACAGAGCTGCTCTTAGAGATTGCTAACTGGAGGAAACACAGAGGACT 257
 Db 124 GTGCGGGGCATGACAGAGCTGCTCTTAGAGATTGCTAACTGGAGGAAACACAGAGGACT 183
 Qy 258 CTCAAATACCAAGTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 317
 Db 184 CTCAAATACCAAGTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 243
 Qy 318 GGTTCAGTGCAGCCCATCTTAAGGCACTTGTCTGGAAGGGCAGAAATGCCAGTGTCTTGGC 377
 Db 244 GGTTCAGTGCAGCCCATCTTAAGGCACTTGTCTGGAAGGGCAGAAATGCCAGTGTCTTGGC 303
 Qy 378 TATGGACCCACAGAGGCTGGAGAGCGCACAAATGCTGGGAGCCAGAGCAACCTGGG 437
 Db 304 TATGGACCCACAGAGGCTGGAGAGCGCACAAATGCTGGGAGCCAGAGCAACCTGGG 363
 Qy 438 GTGATCCCGCGGCTCTCATGGACCTCTGACACTCACAGGAGGAGGAGGCTGCGAGGGC 497
 Db 364 GTGATCCCGCGGCTCTCATGGACCTCTGACACTCACAGGAGGAGGAGGCTGCGAGGGC 423
 Qy 498 CGGCCATGGGCTTCTGTCACCATGCTTACTAGAGATCTACAGAGAGAGTATTA 557
 Db 424 CGGCCATGGGCTTCTGTCACCATGCTTACTAGAGATCTACAGAGAGAGTATTA 483
 Qy 558 GACCTCTTGAGACCTCTTCCGAGAGCTTGTAACTCCGAGAGCTGCGGGGGATATC 617
 Db 484 GACCTCTTGAGACCTCTTCCGAGAGCTTGTAACTCCGAGAGCTGCGGGGGATATC 543
 Qy 618 CTGATTCGGGCTCTCTCCAGAGCCATCAGTAGCTTGTGCTGATTTTGACGCGCACTTC 677
 Db 544 CTGATTCGGGCTCTCTCCAGAGCCATCAGTAGCTTGTGCTGATTTTGACGCGCACTTC 603
 Qy 678 CTGCCAGCAGTGGAAATCGAGCTGAGAGCCACCCGGCTCAACAGAGCTCTCTCCGC 737
 Db 604 CTGCCAGCAGTGGAAATCGAGCTGAGAGCCACCCGGCTCAACAGAGCTCTCTCCGC 663
 Qy 738 AGTCATGCTGCTGCTCTGCTCAAGTGGAGCCAGCGGGAAGTGTGGCCCATTTCCGCG 797
 Db 664 AGTCATGCTGCTGCTCTGCTCAAGTGGAGCCAGCGGGAAGTGTGGCCCATTTCCGCG 723
 Qy 798 CGAGAGGAAAACTCTACCTGATTGACTTTGGCTGGGTCAAGAGCAACCGCGCACAGGC 857
 Db 724 CGAGAGGAAAACTCTACCTGATTGCTTGGCTGGGTCAAGAGCAACCGCGCACAGGC 783
 Qy 858 AACAGGGCTTCCGGCTAAAGAGAGTGGAGCCATCAACCTCTCTGTTTGTCTGGGC 917
 Db 784 AACAGGGCTTCCGGCTAAAGAGAGTGGAGCCATCAACCTCTCTGTTTGTCTGGGC 843
 Qy 918 AAAGTGGTAGATGCGCTGAATCAGGGCTCTCCCTGCTGTAACCTTATCGGACAGCAAGCTC 977
 Db 844 AAAGTGGTAGATGCGCTGAATCAGGGCTCTCCCTGCTGTAACCTTATCGGACAGCAAGCTC 903
 Qy 978 ACTCGCTATTGCAAGACTCTCTGGGTGGCTCAGCCCAAGTATCTTATTTGCAACATT 1037
 Db 904 ACTCGCTATTGCAAGACTCTCTGGGTGGCTCAGCCCAAGTATCTTATTTGCAACATT 963
 Qy 1038 GCCCTGTGAGAGGCTTCTACCTAGACAGTCTCGGCACTCAACTTCTGTCAGAGTCC 1097
 Db 964 GCCCTGTGAGAGGCTTCTACCTAGACAGTCTCGGCACTCAACTTCTGTCAGAGTCC 1023


```

QY 1098 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCCTGCAGCCTCATGCTTGGACCT 1157
Db 1024 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCCTGCAGCCTCATGCTTGGACCT 1083
QY 1158 GTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGAGCCCGAGCCCT 1217
Db 1084 GTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGAGCCCGAGCCCT 1143
QY 1218 GAGGAAGAGAGATTTGGGAGCCCTCAGCCCATGAGCAGCTTCCAGCCTCTGCTCCAGAAA 1277
Db 1144 GAGGAAGAGAGATTTGGGAGCCCTCAGCCCATGAGCAGCTTCCAGCCTCTGCTCCAGAAA 1203
QY 1278 CTCAGCCCTCAGAAAGCTAAGAGCATAGAGCATGAGCCCGCCCATGCTGGAGCCCTCTCAGC 1337
Db 1204 CTCAGCCCTCAGAAAGCTAAGAGCATGAGCATGAGCCCGCCCATGCTGGAGCCCTCTCAGC 1263
QY 1338 TTGGACCGTCTCTGCTTGGCTCCAGGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1397
Db 1264 TTGGACCGTCTCTGCTTGGCTCCAGGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1323
QY 1398 CGAGAGCGGATGCTGCTAATGAAGACAGTAGAAGAGAGAGCCTAGAGATTGAGAGGCTT 1457
Db 1324 CGAGAGCGGATGCTGCTAATGAAGACAGTAGAAGAGAGAGCCTAGAGATTGAGAGGCTT 1383
QY 1458 AAGACGAAGCAAAAGAACTGGAGGCCAAGATGTTGGCCAGAGAGGCTGAGGAAAAGGAG 1517
Db 1384 AAGACGAAGCAAAAGAACTGGAGGCCAAGATGTTGGCCAGAGAGGCTGAGGAAAAGGAG 1443
QY 1518 AACCATTTGCCACATGTGA 1538
Db 1444 AACCATTTGCCACATGTGA 1464

```

RESULT 5

```

US-10-093-317-7
; Sequence 7, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093,317
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Human
US-10-093-317-7

```

```

Query Match 95.0%; Score 1461; DB 4; Length 1464;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 GGTCTGCTCGGCTAAGCAAGATTGGAGCTACTCTGTCTCAACCTCCAGCTCGCGTAAGG 137
Db 4 GGTCTGCTCGGCTAAGCAAGATTGGAGCTACTCTGTCTCAACCTCCAGCTCGCGTAAGG 63

QY 138 GTGGCTGTGGCTCGGCTAAGCAAGATTGGAGCTACTCTGTCTCAACCTCCAGCTCGCGTAAGG 197
Db 64 GTGGCTGTGGCTCGGCTAAGCAAGATTGGAGCTACTCTGTCTCAACCTCCAGCTCGCGTAAGG 123

QY 198 GTGGGGGCTGGAAGCAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAACCAAGAGAGACT 257
Db 124 GTGGGGGCTGGAAGCAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAACCAAGAGAGACT 183

QY 258 CTCAAATACCAAGTTTGTGCTCTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 317
Db 184 CTCAAATACCAAGTTTGTGCTCTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 243

```

```

QY 318 GGTTCAGTGCAGCCCATCTTAAGGCACTTGTGGAAGGCAAGATGSCAGTGTGCTTGGC 377
Db 244 GGTTCAGTGCAGCCCATCTTAAGGCACTTGTGGAAGGCAAGATGSCAGTGTGCTTGGC 303
QY 378 TATGAGCCCAAGAGAGCTGGGAAGACGACACAATGCTGGGAGCCCAAGAGCAACCTGGG 437
Db 304 TATGAGCCCAAGAGAGCTGGGAAGACGACACAATGCTGGGAGCCCAAGAGCAACCTGGG 363
QY 438 GTGATCCCGGGCTCTCATGAGACCTCTGAGCTCTCAAGGGAGAGGGTCCGAGGGC 497
Db 364 GTGATCCCGGGCTCTCATGAGACCTCTGAGCTCTCAAGGGAGAGGGTCCGAGGGC 423
QY 498 CGGCAATGGGCTCTTCTGTCAACATGCTTACCTAGAGATCTACAGAGAGAGGTATTATTA 557
Db 424 CGGCAATGGGCTCTTCTGTCAACATGCTTACCTAGAGATCTACAGAGAGAGGTATTATTA 483
QY 558 GACCTCTGAGACCTCTGCTGGGAGACCTGTAATCCGAGAGACTTGGCGGGGAATATTC 617
Db 484 GACCTCTGAGACCTCTGCTGGGAGACCTGTAATCCGAGAGACTTGGCGGGGAATATTC 543
QY 618 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTTGAGCGGCACTTC 677
Db 544 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTTGAGCGGCACTTC 603
QY 678 CTGCGAGCAGTTCGAAATCGGACTGTAGAGCCACCCGGCTCAACAGAGGCTCTCCCGC 737
Db 604 CTGCGAGCAGTTCGAAATCGGACTGTAGAGCCACCCGGCTCAACAGAGGCTCTCCCGC 663
QY 738 AGTCATGCTGTCTCTGCTCAAGGTGGAACGCGGGAAACGTTTGGCCCCCATTTCCGCAAG 797
Db 664 AGTCATGCTGTCTCTGCTCAAGGTGGAACGCGGGAAACGTTTGGCCCCCATTTCCGCAAG 723
QY 798 CGAGAGGAGAAATCTACTGATTTGACTTGGCTGAGGCAACCGGGGCAAGGC 857
Db 724 CGAGAGGAGAAATCTACTGATTTGACTTGGCTGAGGCAACCGGGGCAAGGC 783
QY 858 AACAGGGCTCTCGGCTAAAGAGAGTGGAGCATCAACACCTCTGTTTGTCTCTGGGC 917
Db 784 AACAGGGCTCTCGGCTAAAGAGAGTGGAGCATCAACACCTCTGTTTGTCTCTGGGC 843
QY 918 AAAGTGTGATGCTGCTGAATCAGGGCTCTCTGCTGATACCTTATCGGAGACAGAGCTC 977
Db 844 AAAGTGTGATGCTGCTGAATCAGGGCTCTCTGCTGATACCTTATCGGAGACAGAGCTC 903
QY 978 ACTGCTCTATTGCAAGACTCTCTGGTGGCTCAGGCCACAGTATCTTATGCGCAACATT 1037
Db 904 ACTGCTCTATTGCAAGACTCTCTGGTGGCTCAGGCCACAGTATCTTATGCGCAACATT 963
QY 1038 GCCCTGAGAGACGCTTCTACTAGACACAGTCTCCGCACTCAACTTTGCTGCGCAGGTCC 1097
Db 964 GCCCTGAGAGACGCTTCTACTAGACACAGTCTCCGCACTCAACTTTGCTGCGCAGGTCC 1023
QY 1098 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCCTGCAGCCTCATGCTTGGGACCT 1157
Db 1024 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCCTGCAGCCTCATGCTTGGGACCT 1083
QY 1158 GTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGAGCCCGAGCCCT 1217
Db 1084 GTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGAGCCCGAGCCCT 1143
QY 1218 GAGGAAGAGAGATTTGGGAGCCCTCAGCCCATGAGCAGCTTCCAGCTCTGCTCCAGAAA 1277
Db 1144 GAGGAAGAGAGATTTGGGAGCCCTCAGCCCATGAGCAGCTTCCAGCTCTGCTCCAGAAA 1203
QY 1278 CTCAGCCCTCAGAAAGCTAAGAGCATGAGCAGCCCGCCCATGCTGGAGCGCTCTCTCAGC 1337
Db 1204 CTCAGCCCTCAGAAAGCTAAGAGCATGAGCAGCCCGCCCATGCTGGAGCGCTCTCTCAGC 1263
QY 1338 TTGAGACGCTGTGCTTGGCTCCAGGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1397
Db 1264 TTGAGACGCTGTGCTTGGCTCCAGGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1323

```



```

; OTHER INFORMATION: protein gene Hskip3a (Figure 1).
; OTHER INFORMATION: Description of Artificial Sequence: Hskip3a gene.
US-09-883-096-1

Query Match          9.0%; Score 138.6; DB 4; Length 4108;
Best Local Similarity 51.0%; Pred. No. 3.8e-29;
Matches 448; Conservative 0; Mismatches 404; Indels 27; Gaps 4;

Qy 270 TTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCAAGTTTCAGTGCAG 329
    |||||
Db 344 TTTGACGGGTCTTTGGCGAGCGCGCCACCAACAGAGCGTGTCCAGCACACACGAC 403

Qy 330 CCCATCTTAAGCACTTCTGTAAGGAGGAGCAATGCCAGTGTCTTGTCTATGACCCACA 389
    |||||
Db 404 AGCGTCTCGGACAGCTTCTCCAGGGTACAACTGCTCAGTGTCTTGTGCTTACGGGCGCAC 463

Qy 390 GGAGCTGGGAGAGCAGACACATGCTGGGCGAGCCAGCACTGGGGTGTATCCCGGG 449
    |||||
Db 464 GGGGCTGGGAGAGACACACCATGCTGGGAAGGGAGGGGACCCCGGCATCATGTATACCTG 523

Qy 450 GCTCTCATGACCTCTCTGAGCTCAAGGAGGAGGAGGTGCGGAGCGCGCCATGGGCC 509
    |||||
Db 524 AC-----CACGTGGAAGTGTACAGGCGCTGGAGCGCGCCAGCAGAGAGAC 574

Qy 510 CTTTCTGTCAATGCTTACTAGATCTACAGAGAGAGGTATTAAGACTCTCTGGAC 569
    |||||
Db 575 TTCGAGGTGCTCATCAGCTACAGAGGTGTATATGAACAGATCCATGACTCTCTGGAG 634

Qy 570 CTTGCTTCGGGAGACCTGTTATCCGAGAGACTCCCGGGGATATCTGATTCGGGT 629
    |||||
Db 635 C---CCAAAGGGGCGCTTGGCCATCCGCGAGACCCCGACAAAGGGGGTGTGTGCAAGGA 691

Qy 630 CTCTCCAGAGAGCCCATCAGTAGCTTGTCTGATTTGTAGCGGCACTTCTCTGCGACCAAGT 689
    |||||
Db 692 CTTTCTTCCACAGCAGCTCAGCCAGCAGCTGTCTGGAGATCTACACAGGGGGAGAC 751

Qy 690 GCAATCGGACTGTAGAGACCAACCGGCTCAACAGAGCTCTCTCCGAGTCTATCTGTG 749
    |||||
Db 752 GGTAAACGCGCAGCAGCAGCCCACTGATGCCAAGCGGACTTCTCTCGGCTCCCATGCCATC 811

Qy 750 CTCCTGGTTCAGGTGACAGGGGAGCGTTGGGCCCATTTTGGCCAG-----CGAGAG 803
    |||||
Db 812 TTCAGATCTTTTGTGAAGCAGCAGGACCGGGTTCAGGACTGACCCAGGCTGTCTCAGGTG 871

Qy 804 GGAATACTCTACCTGATTGACTTGGCTGGGTGAGAGCAACCGCGCAGCAGGCAACAAAG 863
    |||||
Db 872 GCCAAGATGAGCTGATTGACTTGGCTGGGTGAGAGCGGGCATCCAGCACCCATGCCAG 931

Qy 864 GGCCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCTCTCTTGTCTGGGCAAGTG 923
    |||||
Db 932 GGGGAGCGGCTGCGGGAGGGGGCCAACTCAACCGCTCTCTGCTGGGCTCATCAACGTC 991

Qy 924 GTAGATGGCTG-----AATCAGGGCTCTCTGCTGTACTTATTCGGGACAGCAG 974
    |||||
Db 992 CTCATGCTTTGGCGGATGCAAGGGCGCGAAGACCCATGTGCTCCCTACCGGACAGCAAA 1051

Qy 975 CTCACCTCCCTATTGACAGACTCTCTGGGTGCTCAGGCCACAGTATCTTATTGGCCAA 1034
    |||||
Db 1052 CTGACCCGCTCTCAAGAGATCTCTCGGGGGAAGTCTCCGACAGTATGATGCTGCTGCC 1111

Qy 1035 ATTGCCCCCTGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAG 1094
    |||||
Db 1112 ATCAGCCCTCCAGCTGACCTACGAGGACAGGTACAAACACCCCTCAATATGCCGACCGG 1171

Qy 1095 TCCAAGGAGGTGATCAATCGGCTTTTACCAATGAGAGC 1133
    |||||
Db 1172 GCCAAGGAGATCAGGCTCTCTGCTGAAGAGCAATGTGACC 1210

```

RESULT 11
 US-09-883-096-6
 ; Sequence 6, Application US/09883096
 ; Patent No. 6680369

```

; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Craven, Andrew
; APPLICANT: Yu, Ming
; APPLICANT: Sakowicz, Roman
; APPLICANT: Patel, Umesh A.
; APPLICANT: Davies, Katherine A.
; TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE
; FILE REFERENCE: 020552-001410US
; CURRENT APPLICATION NUMBER: US/09/883,096
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 09/594,655
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1152
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Hskip3a
; OTHER INFORMATION: fragment
; OTHER INFORMATION: Nucleotide sequence of Hskip3a fragment used in
; OTHER INFORMATION: Atpase assay (Figure 5).
US-09-883-096-6

Query Match          8.9%; Score 136.6; DB 4; Length 1152;
Best Local Similarity 51.2%; Pred. No. 3.5e-29;
Matches 437; Conservative 0; Mismatches 389; Indels 27; Gaps 4;

Qy 270 TTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGAGCATCTATGCAAGTTTCAGTGCAG 329
    |||||
Db 214 TTTGACCGGGTCTTTGGCGAGCGCGCCACCAACAGAGCGTGTTCACGACACACCGCAC 273

Qy 330 CCCATCTTAAGCACTTCTGGAAGGGCAGAAATGCCAGTGTCTTGTCTATGACCCACA 389
    |||||
Db 274 AGCGTCTCGGACAGCTTCTCTCAGGGCTACAACCTGCTCAGTGTCTTGTGCTTACGGGCGCAC 333

Qy 390 GGAGCTGGGAGAGCAGCACAAATCTGGGCGAGCCAGAGCAACCTGGGGTGTATCCCGGG 449
    |||||
Db 334 GGGGCTGGGAGAGACACACACCATCTCTGGGAAGGAGGGGAGCCCGGCATCATGTA---- 389

Qy 450 GCTCTCATGAGACTCTCTGAGCTCAAGAGGAGAGGGTGCAGAGGGCGCGCCATGGGCC 509
    |||||
Db 390 -----CTGTACACCGTGGAACTGTACAGGCGCTGGAGCGCGCGCAGCAGAGAGAC 444

Qy 510 CTTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGAGGTATTAGACTCTCTGAC 569
    |||||
Db 445 TTCGAGTGTCTCATCAGCTACCGAGGAGTGTATTAATGAACAGATCCATGACTCTCTGGAG 504

Qy 570 CTTCTTCGGGAGACCTGGTAAATCCGAGAGACTGCGGGGGAATATCTGATTCGGGT 629
    |||||
Db 505 C---CCAAAGGGGCGCTTGGCATCCGAGAGACCCCGAAGGGGGTGTGTGCAAGGA 561

Qy 630 CTCTCCAGAGGCCATCAGTAGCTTTGCTGATTTTGAGCGGCACTTCTCTCCAGCCAGT 689
    |||||
Db 562 CTTTCTTTCCACAGCAGCCCTCAGCGAGCAGCTGTCTGGAGATCTGACCAAGGGGAA 621

Qy 690 CGAATCGGACTGTAGAGAGCCACCGGCTCAACAGAGGCTCTCTCCGAGTATGCTGTG 749
    |||||
Db 622 CGTAAACGCGCAGCAGCAGCCCACTGATGCCAAACGCACTTCTCTCCGCTCCCATGCCATC 681

Qy 750 CTCTGCTGCAAGTGGGAGCCAGCGGAAAGTTTGGCCCATTTTGGCCAG-----CGAGAG 803
    |||||
Db 682 TTCAGATCTTTTGTGAAGCAGCAGCGGTTCCAGGACTGACCCAGGCTGTCCAGGTG 741

Qy 804 GGAATACTCTACCTGATTTGACTTGGCTGGGTGAGAGCAACCGGCGCAGGCAACAAAG 863
    |||||
Db 742 GCCAAGATGAGCTGTGATTGACTGCTGGTGGCTCAGAGCGGGCATCCAGCACCCATGCCAAG 801

Qy 864 GGCCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCTCTCTGCTTGTGCTGGGCAAGTG 923
    |||||
Db 802 GGGGAGCGGCTGCGGGAGGGGGCCAAATCAACCGCTCTCTGCTGCGGCTCATCAACGCTC 861

```



```

Db 226 CCCACAGGCTGTGGGAAACCTACACCATGCTGGGCACAGACCAGGAGCTGGCATCTAT 285
Qy 444 CCGCGGGCTCTCATGGACCTCTGAGCTCAACAGGGAGGAGGTCGCGAGGCGCGCA 503
Db 286 GTTCAGACCTCAAGCACTCTTCGTCGCATCGAGGAGACAGCAATGA-----C 336
Qy 504 TGGGCCCCCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTATTAGACCTC 563
Db 337 ATGAGATATGAGGTCTCCATGCTCTACCTGAGATCTACAATGAGATATCGGGACCTG 396
Qy 564 CTGAGACCTGCTTGGGAGACCTGTAATCCGAGAGAGTCCGCGGGGAATATCTGATTT 623
Db 397 CTGAACCCCTCCCTGGGTACCTGAGCTCGGGAGGACTCTAAGGGGGTGTATCCAGGTG 456
Qy 624 CCGGCTCTCCCAAGCCCATCAGTAGCTTTGCTGATTTTGAGCGGCACCTCTCTGCCA 683
Db 457 GCGGCTATCCGAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 516
Qy 684 GCCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCTCCGCAAGTCA 743
Db 517 GGGAAACCGGAGAGGACCCAGAGCCACGGCGCCCAACAGAGCTCTCCGCTCCAC 576
Qy 744 GCTGTGCTCTGCTCAAGGTGACAGCGGGAACGTTTGGCCCCATTTTCGCCAGGAGAG 803
Db 577 GCGGCGCCCTGTTCA----- 592
Qy 804 GGAATACTCTACCTGATTTGCTGGTTCAGAGGACCAACCGCGGCACAGGCAACAG 863
Db 593 -----TGATCGACTGGCTGGCTCAGAGCGCGCTCGCAGACACAGAATCGT 639
Qy 864 GCGCTTGGCTTAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTGGGCAAAAGTG 923
Db 640 GGGCAGCGTATGAAGGAGGGGCCCAATCAACGCTCACTGTGGGCTGGGCAACTGC 699
Qy 924 GTAGATCGCTGATATCAG-----GGCCTCCCTCGTGTACCTTATCGGGAAGCAAGTCT 977
Db 700 ATCAACGCCCTGAGCGCAAGGGTAGCAACAGTACATCACTATCGCGACAGAGCTC 759
Qy 978 ACTGCGCTATTGAGGAGCTCTCTGGTGGTTCAGCCCAAGTATCTTATTTGGCAACATT 1037
Db 760 ACCGCGCTCTGAGGAGCTCTCTGGGAGGAAACAGCGGCACAGTGTATCGCTCACATC 819
Qy 1038 GCCCTGAGAGCGCTTCTACCTAGACAGCTCTCCGCACTCAACTTTGCTGCCAGGTCC 1097
Db 820 AGTCTCGAGAGTGTGCTTCGAGAGTCCCGGAACACCTGACCTACGCGCGCGGCCC 879
Qy 1098 AAGAGGTGATCAATCGG 1115
Db 880 AAGAACAATTAAGACTAGG 897

```

RESULT 14

```

US-09-724-516-1
; Sequence 1, Application US/09724516
; Patent No. 6391573
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6391573el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1045
; CURRENT APPLICATION NUMBER: US/09/724,516
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/621,233
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 897
; TYPE: DNA
; ORGANISM: Human
; FEATURE:

```

```

; NAME/KEY: misc feature
; LOCATION: (472)...(495)
; OTHER INFORMATION: n = a, c, t, or g
US-09-724-516-1

Query Match      8.4%; Score 128.8; DB 3; Length 897;
Best Local Similarity 50.1%; Pred. No. 5.4e-27;
Matches 430; Conservative 0; Mismatches 356; Indels 72; Gaps 3;

Qy 264 TACCAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATCAGGTTCA 323
Db 106 TACCTGTTTCGAGCTGGGCTTTGACTTACCCGCCAGGAGATGGTGTATCAGGCCACC 165
Qy 324 GTGCAGCCCATCTTAAGGCATCTGCTGAAGGAGCAATGCCAGTGTGCTTGCCTTATGA 383
Db 166 ACCAAGACCTCATCGAGGGGCTCATCTCAGGCTACATGCACTGTCTTTGCCATATGGC 225
Qy 384 CCCACAGAGCTGGGAAGACGACACAATGTGGGACGCCAGAGCAACCTGGGGTGATC 443
Db 226 CCCACAGGCTGTGGGAAACCTACACCATGCTGGGCAAGAGAGGCTGGCATCTAT 285
Qy 444 CCGGGGCTCTCATGGACCTCTCTGAGCTCAAGGGAGGAGGTGCGGAGGCGCGCA 503
Db 286 GTTCAGACCCCTCAACGACCTCTTCCGTGCCATCGAGGAGACCAATGA-----C 336
Qy 504 TGGGCCCCCTTCTGTACCATGCTTACCTAGAGATCTACACAGGAGAGCTATTAGACCTC 563
Db 337 ATGAGATATGAGGTCTCCATGCTCTACCTGAGATCTACAATGAGATATCGGGACCTG 396
Qy 564 CTGAGACCTGCTTGGGAGACCTGCTAATCCGAGAAGCTCGCGGGGAATATCTCTGATT 623
Db 397 CTGAACCCCTCCCTGGGCTACTGAGCTGCGGAGGAGCTCTAAGGGGGTGTATCCAGGTG 456
Qy 624 CCGGCTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTTGTGAGCGGCACCTTCTTGCCA 683
Db 457 GCGGCTATCCGAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 516
Qy 684 GCCAGTCAATAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCTCCGCAAGTCA 743
Db 517 GGGAAACCGGAGAGGACCCAGGAGCCCAACGCGGCACAGAGCTCTCCGCTCCAC 576
Qy 744 GCTGTGCTCTGCTCAAGGTGACAGCGGGAACGTTTGGCCCCATTTTCGCCAGGAGAG 803
Db 577 GCGGCGCCCTGTTCA----- 592
Qy 804 GGAATACTCTACCTGATTTGCTGGTTCAGAGGACCAACCGCGGCACAGGCAACAG 863
Db 593 -----TGATCGACTGGCTGGCTCAGAGCGCGCTCGCAGACACAGAATCGT 639
Qy 864 GCGCTTGGCTTAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTGGGCAAAAGTG 923
Db 640 GGGCAGCGTATGAAGGAGGGGCCCAATCAACGCTCACTGTGGGCTGGGCAACTGC 699
Qy 924 GTAGATCGCTGATATCAG-----GGCCTCCCTCGTGTACCTTATCGGGAAGCAAGTCT 977
Db 700 ATCAACGCCCTGAGCGCAAGGGTAGCAACAGTACATCACTATCGCGACAGAGCTC 759
Qy 978 ACTGCGCTATTGAGGAGCTCTCTGGTGGTTCAGCCCAAGTATCTTATTTGGCAACATT 1037
Db 760 ACCGCGCTCTGAGGAGCTCTCTGGGAGGAAACAGCGGCACAGTGTATCGCTCACATC 819
Qy 1038 GCCCTGAGAGCGCTTCTACCTAGACAGCTCTCCGCACTCAACTTTGCTGCCAGGTCC 1097
Db 820 AGTCTCGAGAGTGTGCTTCGAGAGTCCCGGAACACCTGACCTACGCGCGCGGCCC 879
Qy 1098 AAGAGGTGATCAATCGG 1115
Db 880 AAGAACAATTAAGACTAGG 897

```

RESULT 15

```

US-10-090-695-1
; Sequence 1, Application US/10090695

```


Query Match	100.0%;	Score 1538;	DB 17;	Length 1538;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1538;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

QY	1	ATGCAGCCGGCGGCTCGACCCAGCAGAGGCGACCGCAGATGGCGCGAGCTTCAGCGG	60
Db	1	ATGCAGCCGGCGGCTCGACCGCAGCAGAGGCGACCGCAGATGGCGCGAGCTTCAGCGG	60
QY	61	CGCGCATCTCAGGAGCTGGTCCGTGTCCGCTAAGCAAGATTGGAGCTACTCGTCTGTCAC	120
Db	61	CGCGCATCTCAGGAGCTGGTCCGTGTCCGCTAAGCAAGATTGGAGCTACTCGTCTGTCAC	120

121 CTCCAGCTCGGTAAAGGTCGTGCGACTGCGGCCATTGTGGATGGAACACGGGAG 180
 121 CTCCAGCTCGGTAAAGGTCGTGCGACTGCGGCCATTGTGGATGGAACACGGGAG 180
 181 CAAGTGTATCCCTCCCTGTCGTGCGGGGCAATGACAGCTGCTCTCTAGAGATTGCTAACTGGA 240
 181 CAAGTGTATCCCTCCCTGTCGTGCGGGGCAATGACAGCTGCTCTCTAGAGATTGCTAACTGGA 240
 241 GGAACACACAGGAGACTCTCAAAATACCAATGATGCTTCTATGCGGAGAGGAGTACTC 300
 241 GGAACACACAGGAGACTCTCAAAATACCAATGATGCTTCTATGCGGAGAGGAGTACTC 300
 301 AGCAGGACATCTATGCAAGGTCAGTGACAGCCCTCTTAAGGCACTTCTGCTGGAAGGAGCAGA 360
 301 AGCAGGACATCTATGCAAGGTCAGTGACAGCCCTCTTAAGGCACTTCTGCTGGAAGGAGCAGA 360
 361 ATGCCAGTGTGCTTGGCTATGGAACCCACAGAGCTGGGAGAGCGACACATGCTGGCA 420
 361 ATGCCAGTGTGCTTGGCTATGGAACCCACAGAGCTGGGAGAGCGACACATGCTGGCA 420
 421 GCCCAGAGCAACCTGGGGTATCCCGGGGCTCTCATGACCTCTGACAGCTCAAGGG 480
 421 GCCCAGAGCAACCTGGGGTATCCCGGGGCTCTCATGACCTCTGACAGCTCAAGGG 480
 481 AGGAGGTGCGAGGCGCGCCATGGCCCTTCTGTCAACATGCTTACCTAGAGATCT 540
 481 AGGAGGTGCGAGGCGCGCCATGGCCCTTCTGTCAACATGCTTACCTAGAGATCT 540
 541 ACCAGAGAGAGGTATTAGACTCTCTGACCTCTGCGGAGACCTTGGTAATCCGGAAG 600
 541 ACCAGAGAGAGGTATTAGACTCTCTGACCTCTGCGGAGACCTTGGTAATCCGGAAG 600
 601 ACTGCGGGGGAATATCTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTCCTG 660
 601 ACTGCGGGGGAATATCTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTCCTG 660
 661 ATTTTGAGCGGCACTTCTCCAGCAGTGGAAATCGGACTGTAGAGCCACCGGCTCA 720
 661 ATTTTGAGCGGCACTTCTCCAGCAGTGGAAATCGGACTGTAGAGCCACCGGCTCA 720
 721 ACCAGCGCTCTCCCGCAGTCACTGCTGCTCTGCTCAAGGTGGACCAAGCGGAAAGTT 780
 721 ACCAGCGCTCTCCCGCAGTCACTGCTGCTCTGCTCAAGGTGGACCAAGCGGAAAGTT 780
 781 TGGCCCATTTGCGCAGCAGAGAGGAAATCTTACCTGATTTGCTGGCTGGCTCAGAG 840
 781 TGGCCCATTTGCGCAGCAGAGAGGAAATCTTACCTGATTTGCTGGCTGGCTCAGAG 840
 841 ACAACCGGCGCACAGGCAACAGGGCCCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCT 900
 841 ACAACCGGCGCACAGGCAACAGGGCCCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCT 900
 901 CCCTGTTTGTCTGGGCAAAAGTGTAGATCGCTGTAATCAGGCGCTCCCTCGTGTACCTT 960
 901 CCCTGTTTGTCTGGGCAAAAGTGTAGATCGCTGTAATCAGGCGCTCCCTCGTGTACCTT 960
 961 ATCGGACAGAGACTCACTGCGCTATGAGAGAGCTCTGCGGTGGCTCAGCCACAGTA 1020
 961 ATCGGACAGAGACTCACTGCGCTATGAGAGAGCTCTGCGGTGGCTCAGCCACAGTA 1020
 1021 TCCTTATGCGCAACATTTGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCA 1080
 1021 TCCTTATGCGCAACATTTGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCA 1080
 1081 ACTTGTGTCAGGTCCAGAGAGGTGATCAATCGGCGCTTTTACCAATGAGAGCTCGAGC 1140
 1081 ACTTGTGTCAGGTCCAGAGAGGTGATCAATCGGCGCTTTTACCAATGAGAGCTCGAGC 1140
 1141 CTATGCTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGCA 1200
 1141 CTATGCTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGCA 1200
 1201 AGAGAGCCGAGGCGCTGAGAGAGGAGATTTGGAGCCCTTGAGCCCTGAGCTCCAG 1260

1201 AGAGAGCCCGAGGCTCTGAGAGAGGAGATTGGGAGCCCTGAGCCATGGCAGCTCCAG 1260
 1261 CCTCTGCTCCCAAGAACTCAGCCCCCTTACAGAGCTTAAGCAGCATGGACCCCGCCATGC 1320
 1261 CCTCTGCTCCCAAGAACTCAGCCCCCTTACAGAGCTTAAGCAGCATGGACCCCGCCATGC 1320
 1321 TGGAGCGCTCTCTCAGCTTGGACCGCTCTGCTTGCCTCCAGGGGAGCCAGGGGGCCCTC 1380
 1321 TGGAGCGCTCTCTCAGCTTGGACCGCTCTGCTTGCCTCCAGGGGAGCCAGGGGGCCCTC 1380
 1381 TGTGTAGTATCCCAAGAGCGAGCGGATGGTGTCTAATGAAGACAGTAGAAGAGAAGGACC 1440
 1381 TGTGTAGTATCCCAAGAGCGAGCGGATGGTGTCTAATGAAGACAGTAGAAGAGAAGGACC 1440
 1441 TAGAGATTGAGAGGCTTAAAGCAAGCAAAAAGAACTGGAGGCAAGATGTTGGCCGAGA 1500
 1441 TAGAGATTGAGAGGCTTAAAGCAAGCAAAAAGAACTGGAGGCAAGATGTTGGCCGAGA 1500
 1501 AGGCTGAGAAAGAGGAGAACCATTTGTCACCATGTGA 1538
 1501 AGGCTGAGAAAGAGGAGAACCATTTGTCACCATGTGA 1538

RESULT 2
 US-10-334-143-102
 ; Sequence 102, Application US/10334143
 ; Publication No. US20040009549A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH
 ; APPLICANT: SUDARSANAM, SUCHA
 ; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
 ; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD
 ; FILE REFERENCE: 038602/1543
 ; CURRENT APPLICATION NUMBER: US/10/334,143
 ; PRIOR FILING DATE: 2002-12-31
 ; PRIOR FILING DATE: 2001-12-31
 ; PRIOR FILING DATE: 2001-12-31
 ; NUMBER OF SEQ ID NOS: 207
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 102
 ; LENGTH: 2099
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-334-143-102

Query Match 98.8%; Score 1519; DB 16; Length 2099;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1530; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 6 GCCCGCGCGGCTCGACGCGAGAGGCGGACGCGAGATGGCGGAGCTTCAGCGGCGGCG 65
 DB 29 GCCCGCGCGGCTCGACGCGAGAGGCGGACGCGAGATGGCGGAGCTTCAGCGGCGGCG 88
 QY 66 ATCTCAGGAGCTGGTGGCTGCTGCTAAGCAAGATTGGAGTACTGCTGCTCCACCTCCA 125
 DB 89 ATCTCAGGAGCTGGTGGCTGCTGCTAAGCAAGATTGGAGTACTGCTGCTCCACCTCCA 148
 QY 126 GCTCGGTAAGGCTGGCTGTCGACTGCGGCGCAATTTGTGATGGAACAGCGGAGCAAGT 185
 DB 149 GCTCGGTAAGGCTGGCTGTCGACTGCGGCGCAATTTGTGATGGAACAGCGGAGCAAGT 208
 QY 186 GATCCCCCTGTCGCGGCGGATGGACAGCTGCTCTAGAGATTGCTACTGAGGAAC 245
 DB 209 GATCCCCCTGTCGCGGCGGATGGACAGCTGCTCTAGAGATTGCTACTGAGGAAC 268
 QY 246 CACCAGGAGACTCTCAAAATACCAAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAG 305
 DB 269 CACCAGGAGACTCTCAAAATACCAAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAG 328
 QY 306 GACATCTATGACGTTTCAGTGTGAGCCCATTCCTAAGGCACTTGTGGAAGGCAAGATGCC 365
 DB 329 GACATCTATGACGTTTCAGTGTGAGCCCATTCCTAAGGCACTTGTGGAAGGCAAGATGCC 388

```
Qy 366 AGTGTGCTTGCCTATGACCCACAGAGCTGGGAAGAGCGCACAAATGCTGGGAGGCCA 425
Db 389 AGTGTGCTTGCCTATGACCCACAGAGCTGGGAAGAGCGCACAAATGCTGGGAGGCCA 448
Qy 426 GAGCAACCTGGGGTGATCCCGGGGCTCTCATGACCTCTCGAGCTCACAAGGGAGAG 485
Db 449 GAGCAACCTGGGGTGATCCCGGGGCTCTCATGACCTCTCGAGCTCACAAGGGAGAG 508
Qy 486 GGTGCCGAGGCGCGGCATGCGCCCTTCTGTGCAATGCTCTTACCTAGAGATCTACAG 545
Db 509 GGTGCCGAGGCGCGGCATGCGCCCTTCTGTGCAATGCTCTTACCTAGAGATCTACAG 568
Qy 546 GAGAAGGTATTAGACCTCTCTGACCTCTTCCGAGACCTGGTAATCCGAGAAGACTGC 605
Db 569 GAGAAGGTATTAGACCTCTCTGACCTCTTCCGAGACCTGGTAATCCGAGAAGACTGC 628
Qy 606 CGGGGGAATATCTGATTTCCGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTGATTT 665
Db 629 CGGGGGAATATCTGATTTCCGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTGATTT 688
Qy 666 GAGCGGCACTTCTCGCCAGCCAGTCGAATCGGACTGTAGAGCCACCCGGCTCAACAG 725
Db 689 GAGCGGCACTTCTCGCCAGCCAGTCGAATCGGACTGTAGAGCCACCCGGCTCAACAG 748
Qy 726 CGCTCCTCCCGAGTCATGCTGTCTCTGCTCAAGGTGAGCCAGCGGGAACGTTTGCC 785
Db 749 CGCTCCTCCCGAGTCATGCTGTCTCTGCTCAAGGTGAGCCAGCGGGAACGTTTGCC 808
Qy 786 CCATTTGCCAGCGAGAGGGAATACTCTACCTGATTTGACTTTGGCTGGGTGAGAGCAAC 845
Db 809 CCATTTGCCAGCGAGAGGGAATACTCTACCTGATTTGACTTTGGCTGGGTGAGAGCAAC 868
Qy 846 CGGCGCAGAGCAACAGGGCTTCGGCTAAAGAGGTGAGCGGCATCAACACCTCCCTG 905
Db 869 CGGCGCAGAGCAACAGGGCTTCGGCTAAAGAGGTGAGCGGCATCAACACCTCCCTG 928
Qy 906 TTTGTCTCGGCAAGTGGTAGATGCGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGG 965
Db 929 TTTGTCTCGGCAAGTGGTAGATGCGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGG 988
Qy 966 GACAGCAAGCTCACTCGCTATTGAGAGCTCTCTGGGTGGCTCAGGCCACAGTATCCTT 1025
Db 989 GACAGCAAGCTCACTCGCTATTGAGAGCTCTCTGGGTGGCTCAGGCCACAGTATCCTT 1048
Qy 1026 ATTGCCAATGTCCTCGAGAGCGCTTCTACCTAGACAGCTCCGCACTCAACTTT 1085
Db 1049 ATTGCCAATGTCCTCGAGAGCGCTTCTACCTAGACAGCTCCGCACTCAACTTT 1108
Qy 1086 GCTGCCAGGTCACAGGAGGTGATCAATCGGCTTTTACCATGAGAGCTTCGACGCTCAT 1145
Db 1109 GCTGCCAGGTCACAGGAGGTGATCAATCGGCTTTTACCATGAGAGCTTCGACGCTCAT 1168
Qy 1146 GCCTTGGGACCTGTTAAGCTCTCTCAGAAAGATTGCTTGTGTCACAGAGGCAAGAGA 1205
Db 1169 GCCTTGGGACCTGTTAAGCTCTCTCAGAAAGATTGCTTGTGTCACAGAGGCAAGAGA 1228
Qy 1206 GCGCGAGGCTTGAGGAGAGGATTTGGGAGCCCTGAGCCCATGGCAGCTCCAGCCTCT 1265
Db 1229 GCGCGAGGCTTGAGGAGAGGATTTGGGAGCCCTGAGCCCATGGCAGCTCCAGCCTCT 1288
Qy 1266 GCCTCCAGAAACTCAGCCCTCTACAGAAAGCTAAGAGCAGCATGGACCCGGCCATGCTGGAG 1325
Db 1289 GCCTCCAGAAACTCAGCCCTCTACAGAAAGCTAAGAGCAGCATGGACCCGGCCATGCTGGAG 1348
Qy 1326 CGCTCTCTCAGCTTGGACCGTCTGCTTGCCTCCAGGGAGCCAGGGGGCCCCCTCTGTG 1385
Db 1349 CGCTCTCTCAGCTTGGACCGTCTGCTTGCCTCCAGGGAGCCAGGGGGCCCCCTCTGTG 1408
Qy 1386 AGTACCCCAAGCCAGAGCGGATGGTGTCTATGAGACAGTAGAGAGAGGAGCTTAG-A 1444
Db 1409 AGTACCCCAAGCCAGAGCGGATGGTGTCTATGAGACAGTAGAGAGAGGAGCTTAGTA 1468
```

```
Qy 1445 GATTGAGAGGCTTAAGACGCAAGCAAAAAGAACTGGAGGCCAAGATGTTGGCCAGAGGC 1504
Db 1469 GATTGAGAGGCTTAAGACGCAAGCAAAAAGAACTGGAGGCCAAGATGTTGGCCAGAGGC 1528
Qy 1505 TGAGAAAAGGAGAACCATTTGTCACCAATG 1535
Db 1529 TGAGAAAAGGAGAACCATTTGTCACCAATG 1559

RESULT 3
US-10-797-893-7
; Sequence 7, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10797,893
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Human
US-10-797-893-7

Query Match 95.0%; Score 1461; DB 17; Length 1464;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 78 GGTGCTGTGCGCTTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGGTAAAG 137
Db 4 GGTGCTGTGCGCTTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGGTAAAG 63
Qy 138 GTGCTGTGCGACTGCGGCCCATTTTGTGGATGGAAACAGCGGAGCAAGTGATGCCCTGT 197
Db 64 GTGCTGTGCGACTGCGGCCCATTTTGTGGATGGAAACAGCGGAGCAAGTGATGCCCTGT 123
Qy 198 GTGCGGCGCATGGACAGCTGCTCTCTAGAGATTCTTAACCTGGAGGAAACCCAGAGACT 257
Db 124 GTGCGGCGCATGGACAGCTGCTCTCTAGAGATTCTTAACCTGGAGGAAACCCAGAGACT 183
Qy 258 CTCAAATACCAATTTGATGCTTCTATGGGAGAGGAGTACTCTAGAGGACATCTATGCA 317
Db 184 CTCAAATACCAATTTGATGCTTCTATGGGAGAGGAGTACTCTAGAGGACATCTATGCA 243
Qy 318 GGTTCAGTGCAGGCCCATCTTAAGGCACTTGTGTGAAGGCGCAGAAATGCCAGTGTGCTGCC 377
Db 244 GGTTCAGTGCAGGCCCATCTTAAGGCACTTGTGTGAAGGCGCAGAAATGCCAGTGTGCTGCC 303
Qy 378 TATGGACCCACAGAGCTGGGAGAGACCCACACATGCTGGGACGCCAGCAACCTGGG 437
Db 304 TATGGACCCACAGAGCTGGGAGAGACCCACACATGCTGGGACGCCAGCAACCTGGG 363
Qy 438 GTGATCCCGGGGCTCTCTGAGCCTCTGAGCTCAAGGGAGAGGGGTGCCGAGGGC 497
Db 364 GTGATCCCGGGGCTCTCTGAGCCTCTGAGCTCAAGGGAGAGGGGTGCCGAGGGC 423
Qy 498 CGGCGCATGGGCGCTTCTGTGCAACATGCTTAACTAGAGATCTTACAGGAGAGGTATTA 557
Db 424 CGGCGCATGGGCGCTTCTGTGCAACATGCTTAACTAGAGATCTTACAGGAGAGGTATTA 483
Qy 558 GACCTCTCGACCGCTGCTCGGAGAGCTGCTGTAATCCGAGAGACTCCCGGGGGAATATC 617
Db 484 GACCTCTCGACCGCTGCTCGGAGAGCTGCTGTAATCCGAGAGACTCCCGGGGGAATATC 543
Qy 618 CTGATTCGGGCTCTCTCCAGAGAGCCCATCAGTAGCTTTGCTGATTTTGGCGGCATCTC 677
```

[illegible]

RESULT 4

US-09-925-300-288

US-05-323-300-288
; Sequence 288, Application US/09925300

; sequence zoo, Application
; Patent No. US20020151681A1

; GENERAL INFORMATION:

APPLICANT: Craig Rosen,

APPLICANT: Steve Ruben

TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA101

```

; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 288
; LENGTH: 2104
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (497)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1323)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-288

```

Query Match	94.3%	Score 1450.2;	DB 9;	Length 2104;
Best Local Similarity	99.4%;	Pred. No. 0;		
Matches 1457; Conservative	7;	Mismatches	1;	Indels 1;

Qy	71	AGGAGCTGGTGCCTGTGCGCTAAACAAGATTGGAGCTACTCGTCTGTCACCTCCAGCTCG	130
Db	60	AGGAGCTGGTGCCTGTGCGCTAAACAAGATTGGAGCTACTCGTCTGTCACCTCCAGCTCG	119
Qy	131	CGTAAGGGTGGCTGTGCGACTGCGGCCATTGTGATGGAACAGCGGAGCAAGTGATCC	190
Db	120	CGTAAGGGTGGCTGTGCGACTGCGGCCATTGTGATGGAACAGCGGAGCAAGTGATCC	179
Qy	191	CCCTGTGTGCGGGGCATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGAAACCA	250
Db	180	CCCCTGTGTGCGGGGCATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGAAACCA	239
Qy	251	GGAGACTCTCAAAATACCAGTTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACAT	310
Db	240	GGAGACTCTCAAAATACCAGTTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACAT	299
Qy	311	CTATGCAAGTTTCAGTGCAGGCCCATCTCTAAAGGCACTTGCTGGAAGGGCAGAATGCCAGTGT	370
Db	300	CTATGCAAGTTTCAGTGCAGGCCCATCTCTAAAGGCACTTGCTGGAAGGGCAGAATGCCAGTGT	359
Qy	371	GCTTGGCTATGGACCCACAGAGAGTGGGAGACGCAACAATGCTGGCAGGCCACAGAGCA	430
Db	360	GCTTGGCTATGGACCCACAGAGAGTGGGAGACGCAACAATGCTGGCAGGCCACAGAGCA	419
Qy	431	ACCTGGGGTCATCCCGGGGCTCTCATGGAACCTCTGCAGCTCAACGGGAGGAGGGTGC	490
Db	420	ACCTGGGGTGATCCCGGGGCTCTCATGGAACCTCTGCAGCTCAACGGGAGGAGGGTGC	479
Qy	491	CGAGGGCCGCCATGGGCCCTTTCTGTCAACATGCTCTTACCTAGAGATCTACAGAGAA	550
Db	480	CGAGGGCCGCCATGGGCCCTTTCTGTCAACATGCTCTTACCTAGAGATCTACAGAGAA	539
Qy	551	GGTATTAGACTCCTTGAGACCTGTTCCGGAGACCTGGTAATCCGAGAAGCTGCCGGG	610
Db	540	GGTATTAGACTCCTTGAGACCTGTTCCGGAGACCTGGTAATCCGAGAAGCTGCCGGG	599
Qy	611	GAATATCTCTGATTCCGGGGTCTCTCCAGAAGCCCATCAGTAGCTTTGCTGATTTTGACG	670
Db	600	GAATATCTCTGATTCCGGGGTCTCTCCAGAAGCCCATCAGTAGCTTTGCTGATTTTGACG	659
Qy	671	GCACTTCTGCCAGCCAGTGGAAATCGGACTGTAGGAGCCACCCGGCTCAACCGCGCTC	730
Db	660	GCACTTCTGCCAGCCAGTGGAAATCGGACTGTAGGAGCCACCCGGCTCAACCGCGCTC	719
Qy	731	CTCCCGCAGTCTAGTGTGCTCTCTGGTCAAGGTGGACAGCGGAAACGTTTGGCCCCCAT	790

```

Db 720 CTCGCGAGTCATGCTGTGCTCTCTGCTCAAGGTGGACGAGGGAAAGCTTTGGCCCAATT 779
Qy 791 TCGCAGCGAGAGGAAACTCTACTGATTTGACTTTGGCTGCGGTGAGAGGACAAACGGCG 850
Db 780 TCGCAGCGAGAGGAAACTCTACTGATTTGACTTTGGCTGCGGTGAGAGGACAAACGGCG 839
Qy 851 CACAGGCAACAGGGCCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCCCTGTTTGT 910
Db 840 CACAGGCAACAGGGCCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCCCTGTTTGT 899
Qy 911 CCTGGGCAAGTGTAGATGGCTGAATCAGGCGCTCCCTCTGTGTACCTTATCGGGAAG 970
Db 900 CCTGGGCAAGTGTAGATGGCTGAATCAGGCGCTCCCTCTGTGTACCTTATCGGGAAG 959
Qy 971 CAAGCTCACTCGCTTATTTGAGGAGCTCTCTGGGTGGCTCAGCCACATGATCTTATTCG 1030
Db 960 CAAGCTCACTCGCTTATTTGAGGAGCTCTCTGGGTGGCTCAGCCACATGATCTTATTCG 1019
Qy 1031 CAACATTTGCCCTGAGAGAGCTTCTACTAGACACAGTCTCCGACCTCAACTTTGCTGC 1090
Db 1020 CAACATTTGCCCTGAGAGAGCTTCTACTAGACACAGTCTCCGACCTCAACTTTGCTGC 1079
Qy 1091 CAGTTCGAAGAGTGTATCAATCGGCTTTTACCAATGAGAGCTCTGACGCTCATGCGCTT 1150
Db 1080 CAGTTCGAAGAGTGTATCAATCGGCTTTTACCAATGAGAGCTCTGACGCTCATGCGCTT 1139
Qy 1151 GGGACCTGTTAAGCTCTCAGAAAGATTCCTTGGTCCACAGAGGCAAGAGAGCGCG 1210
Db 1140 GGGACCTGTTAAGCTCTCAGAAAGATTCCTTGGTCCACAGAGGCAAGAGAGCGCG 1199
Qy 1211 AGGCGCTCAGGAAGAGAGATTTGGAGCCCTTGAGCCCATGGCAGCTCTGCGCTC 1270
Db 1200 AGGCGCTCAGGAAGAGAGATTTGGAGCCCTTGAGCCCATGGCAGCTCTGCGCTC 1259
Qy 1271 CCAGAACTAGCCCCCTACAGAGCTTAAGCAGCATGGAACCGGCGCATGCTGGAGCGCT 1330
Db 1260 CCAGAACTAGCCCCCTACAGAGCTTAAGCAGCATGGAACCGGCGCATGCTGGAGCGCT 1319
Qy 1331 CCT-CAGCTTCGACGCTCTGCTGCTCCAGGGGAGCGAGGGGCGCTCTGTGAGTA 1389
Db 1320 CCTTCAGCTTCGACGCTCTGCTGCTCCAGGGGAGCGAGGGGCGCTCTGTGAGTA 1379
Qy 1390 CCCAAAGCAGAGCGAGTGTGTCTAATGAAGACAGTGAAGAGAGAGGACCTAGAGATTG 1449
Db 1380 CCCAAAGCAGAGCGAGTGTGTCTAATGAAGACAGTGAAGAGAGAGGACCTAGAGATTG 1439
Qy 1450 AGAGCTTAAAGCAAGCAAAAGAACTGGAGGCAAGATTTGGCCCAAGGCTGAGG 1509
Db 1440 ARAGCTTAARACGARGCAAAAGAACTGGAGGCAAGATTTGGCCCAAGGCTGAGG 1499
Qy 1510 AAAAGGAGAACCATTTGCCCAATG 1535
Db 1500 AAAAGGAGAACCATTTGCCCAATG 1525

```

RESULT 5

```

; Sequence 1, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/797,893
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 1
; LENGTH: 1115
; TYPE: DNA
; ORGANISM: Human
US-10-797-893-1

Query Match      72.3%; Score 1112.4; DB 17; Length 1115;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1113; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGACAGCGCGGGGGCTCGACGACAGAGCGGCGAGATGCGCGGAGCTTCAGCGG 60
Db 1 ATGACAGCGCGGGGGCTCGACGACAGAGCGGCGAGATGCGCGGAGCTTCAGCGG 60
Qy 61 CGGCGATCTCAGGAGCTGCTGCTGCTAAAGCAAGATTGGAGCTACTCGTCTGCTCAAC 120
Db 61 CGGCGATCTCAGGAGCTGCTGCTGCTAAAGCAAGATTGGAGCTACTCGTCTGCTCAAC 120
Qy 121 CTCCAGCTCGCGTAAAGGCTGCTGTCGACCTGCGGCCAATTTGTGGATGGAACAGCGGAG 180
Db 121 CTCCAGCTCGCGTAAAGGCTGCTGTCGACCTGCGGCCAATTTGTGGATGGAACAGCGGAG 180
Qy 181 CAAGTGATCCCGCTGCTGTCGGGGCATGGACAGCTGCTCTAGAGATTGCTAACTGGA 240
Db 181 CAAGTGATCCCGCTGCTGTCGGGGCATGGACAGCTGCTCTAGAGATTGCTAACTGGA 240
Qy 241 GGAACCAACAGGAGACTCTCAAAATACCAAGTTTGGATGGCTCTTATGGGAGAGGAGTACTC 300
Db 241 GGAACCAACAGGAGACTCTCAAAATACCAAGTTTGGATGGCTCTTATGGGAGAGGAGTACTC 300
Qy 301 AGCAGGACATCTATGCAAGTTTCAAGTTCAGTGCAGCCCATCTTAAGGCACCTTGTGGAAGGGCAGA 360
Db 301 AGCAGGACATCTATGCAAGTTTCAAGTTCAGTGCAGCCCATCTTAAGGCACCTTGTGGAAGGGCAGA 360
Qy 361 ATGCGAGTGTCTTGGCTTATGGAACCAAGAGCTGGGAAAGACGACAACAATGCTGGGCA 420
Db 361 ATGCGAGTGTCTTGGCTTATGGAACCAAGAGCTGGGAAAGACGACAACAATGCTGGGCA 420
Qy 421 GCCCAGAGAACCTGGGGTGTATCCCGGGGTCTCATGAGACCTCTGAGCTCACAAGGG 480
Db 421 GCCCAGAGAACCTGGGGTGTATCCCGGGGTCTCATGAGACCTCTGAGCTCACAAGGG 480
Qy 481 AGGAGGTTGCGAGGGCGCGCATGGGCGCTTCTGTCAACATGCTTCTTACCTAGAGATCT 540
Db 481 AGGAGGTTGCGAGGGCGCGCATGGGCGCTTCTGTCAACATGCTTCTTACCTAGAGATCT 540
Qy 541 ACCAGGAGAGGATTAATAGACCTCTGGAACCTGCTTGGGAGACCTGTGTAATCCGAGAG 600
Db 541 ACCAGGAGAGGATTAATAGACCTCTGGAACCTGCTTGGGAGACCTGTGTAATCCGAGAG 600
Qy 601 ACTCGCGGGGAATATCTGATTCGGGGTCTCTCCAGAACCCATCAGTACTGCTTGTCTG 660
Db 601 ACTCGCGGGGAATATCTGATTCGGGGTCTCTCCAGAACCCATCAGTACTGCTTGTCTG 660
Qy 661 ATTTTGAAGCGGCACTTCTGTCAGCGAGTGAATAATCGGACCTGAGGAGCCACCCGGCTCA 720
Db 661 ATTTTGAAGCGGCACTTCTGTCAGCGAGTGAATAATCGGACCTGAGGAGCCACCCGGCTCA 720
Qy 721 ACCAGCGCTCTCCCGAGTCACTGCTGCTCTGTCCTGTCAGAGTGAGCAGCGGGAACGTT 780
Db 721 ACCAGCGCTCTCCCGAGTCACTGCTGCTCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCT 780
Qy 781 TGGCCCAATTTTCGCGAGGAGGAACTCTACCTGATTTGATTTGCTTGGCTGGGTCAGAGG 840
Db 781 TGGCCCAATTTTCGCGAGGAGGAACTCTACCTGATTTGATTTGCTTGGCTGGGTCAGAGG 840
Qy 841 ACAACCGCGCACAGGCAACAAAGGCGCTTCCGGCTTAAAGAGAGTGGAGCCATCAACACT 900
Db 841 ACAACCGCGCACAGGCAACAAAGGCGCTTCCGGCTTAAAGAGAGTGGAGCCATCAACACT 900
Qy 901 CCCTGTTTGTCTCGGCAAGTGTAGATGCGCTGGAATCAGGCGCTCTCTGCTGTACCTT 960
Db 901 CCCTGTTTGTCTCGGCAAGTGTAGATGCGCTGGAATCAGGCGCTCTCTGCTGTACCTT 960

```


QY 961 ATCGGACAGCAAGCTCACTCGCCTATTGAGGACTCTCTGGGTGGCTCAGCCACAGTA 1020
 Db 961 ATCGGACAGCAAGCTCACTCGCCTATTGAGGACTCTCTGGGTGGCTCAGCCACAGTA 1020
 QY 1021 TCCTTATTGCCAATTTGCCCTTGAGAGAGCGTTCTACCTAGACACAGTCTCCGCACTCA 1080
 Db 1021 TCCTTATTGCCAATTTGCCCTTGAGAGAGCGTTCTACCTAGACACAGTCTCCGCACTCA 1080
 QY 1081 ACTTTGCTGCAGGTCCAGGAGGTGATCAATCG 1114
 Db 1081 ACTTTGCTGCAGGTCCAGGAGGTGATCAATCG 1114

RESULT 6

US-10-797-893-5
 ; Sequence 5, Application US/10797893
 ; Publication No. US20040142397A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; TITLE OF INVENTION: Novel motor proteins and methods for
 ; their use
 ; FILE REFERENCE: 1044
 ; CURRENT APPLICATION NUMBER: US/10/797,893
 ; CURRENT FILING DATE: 2004-03-09
 ; PRIOR FILING DATE: US/09/724,224
 ; PRIOR FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: 09/597,292
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 1041
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-10-797-893-5

Query Match 67.3%; Score 1035.4; DB 17; Length 1041;
 Best Local Similarity 99.9%; Pred. No. 1.5e-306;
 Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 78 GGTGCTGTGCGCTTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG 137
 Db 4 GGTGCTGTGCGCTTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG 63
 QY 138 GTGGCTGTGCGCTTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG 137
 Db 64 GTGGCTGTGCGCTTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG 123
 QY 198 GTGGGGGCAATGACAGCTCTCTAGAGATTGCTAACTGGAGGAAACACACAGGAGACT 257
 Db 124 GTGGGGGCAATGACAGCTCTCTAGAGATTGCTAACTGGAGGAAACACACAGGAGACT 193
 QY 258 CTCAATATACAGTTGATGCTCTTATGGGAGGAGGAGTACTCAGCAGGACATCTATGCA 317
 Db 184 CTCAATATACAGTTGATGCTCTTATGGGAGGAGGAGTACTCAGCAGGACATCTATGCA 243
 QY 318 GGTTCAGTGCAGCCCATCTTAAGGCATCTCTGAGAGGAGGAGGAGTATGCCAGTGTCTTGC 377
 Db 244 GGTTCAGTGCAGCCCATCTTAAGGCATCTCTGAGAGGAGGAGGAGTATGCCAGTGTCTTGC 303
 QY 378 TATGGACCCACAGAGCTGGGAGAGACGACACAAATGCTGGGAGGAGGAGGAGGAGGAGG 437
 Db 304 TATGGACCCACAGAGCTGGGAGAGACGACACAAATGCTGGGAGGAGGAGGAGGAGGAGG 363
 QY 438 GTGATCCCGGGGCTCTCTATGAGACCTCTCTAGGAGGAGGAGGAGGAGGAGGAGGAGG 497
 Db 364 GTGATCCCGGGGCTCTCTATGAGACCTCTCTAGGAGGAGGAGGAGGAGGAGGAGGAGG 423
 QY 498 CGGCCATGGGCGCTTCTGTGTCACCATGCTCTTACCTAGAGATCTACAGGAGGAGGATTA 557
 Db 424 CGGCCATGGGCGCTTCTGTGTCACCATGCTCTTACCTAGAGATCTACAGGAGGAGGATTA 483

QY 558 GACCTCTCGACCCCTCTTCGCGAGACCTGCTGTAATCCGAGAGACTGCGCGGGGAATATC 617
 Db 484 GACCTCTCGACCCCTCTTCGCGAGACCTGCTGTAATCCGAGAGACTGCGCGGGGAATATC 543
 QY 618 CTGATTCGCGGTCTCTCCCAAGAGCCCATCAGTAGCTTTGCTGATTTTGTAGCGGCACTTC 677
 Db 544 CTGATTCGCGGTCTCTCCCAAGAGCCCATCAGTAGCTTTGCTGATTTTGTAGCGGCACTTC 603
 QY 678 CTGCGACAGTGGAAATCGGACTGTAGAGCCACCCGGCTCAACGAGCGCTCTCTCCGC 737
 Db 604 CTGCGACAGTGGAAATCGGACTGTAGAGCCACCCGGCTCAACGAGCGCTCTCTCCGC 663
 QY 738 AGTCATGCTGTCTCTCTGGTCAAGGTGGACCGGGAACGTTTGGCCCCCATTTGCGCAG 797
 Db 664 AGTCATGCTGTCTCTCTGGTCAAGGTGGACCGGGAACGTTTGGCCCCCATTTGCGCAG 723
 QY 798 CGAGAGGGGAAACTCTACCTGATTGATTTGGCTGGGTCAAGAGCAACCGGCGCACAGGC 857
 Db 724 CGAGAGGGGAAACTCTACCTGATTGATTTGGCTGGGTCAAGAGCAACCGGCGCACAGGC 783
 QY 858 AACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTTGGGC 917
 Db 784 AACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTTGGGC 843
 QY 918 AAGTGTGTAGATCGCTGATCAATCAGGGCTCTCTCGTGTACTTTATCGGAGCAGCAAGCTC 977
 Db 844 AAGTGTGTAGATCGCTGATCAATCAGGGCTCTCTCGTGTACTTTATCGGAGCAGCAAGCTC 903
 QY 978 ACTCGCTATTGACAGACTCTCTGGGTGGTCAAGCCACAGTATCTTATTTGCAACATT 1037
 Db 904 ACTCGCTATTGACAGACTCTCTGGGTGGTCAAGCCACAGTATCTTATTTGCAACATT 963
 QY 1038 GCGCCCTGAGAGCGCTTCTACCTAGACAGTCTCGCAGTCAACCTTGTGTCAGGTC 1097
 Db 964 GCGCCCTGAGAGCGCTTCTACCTAGACAGTCTCGCAGTCAACCTTGTGTCAGGTC 1023
 QY 1098 AAGGAGGTGATCAATCG 1114
 Db 1024 AAGGAGGTGATCAATCG 1040

RESULT 7

US-09-960-253-131
 ; Sequence 131, Application US/09960253
 ; Patent No. US20020123619A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Mohamath, Raodoh
 ; APPLICANT: Lodes, Michael J.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; OF LUNG CANCER
 ; FILE REFERENCE: 210121.556
 ; CURRENT APPLICATION NUMBER: US/09/960,253
 ; CURRENT FILING DATE: 2001-09-20
 ; NUMBER OF SEQ ID NOS: 187
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 131
 ; LENGTH: 491
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-960-253-131

Query Match 31.2%; Score 480; DB 9; Length 491;
 Best Local Similarity 99.8%; Pred. No. 2.1e-136;
 Matches 491; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1010 AGCCACAGTATCTTATTGCCAATTCGCGGAGAGCGCTTCTACCTAGACAGT 1069
 Db 1 AGCCACAGTATCTTATTGCCAATTCGCGGAGAGCGCTTCTACCTAGACAGT 60
 QY 1070 CTCGCGACTCAATTTGCTGCGAGGTCGAGGAGGATCAATCGGCTTTTACCAATGA 1129
 Db 61 CTCGCGACTCAATTTGCTGCGAGGTCGAGGAGGATCAATCGGCTTTT-CCAATGA 119

QY 1130 GAGCCTGCAGCCTCATGCTTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTCCTTGGTCC 1189
 Db 120 GAGCCTGCAGCCTCATGCTTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTCCTTGGTCC 179
 QY 1190 ACCGAGCAAGAGAGAGCCGAGGCTTGGAGAGAGAGAGATTGGAGCCCTTGAGCCCAT 1249
 Db 180 ACCGAGCAAGAGAGAGCCGAGGCTTGGAGAGAGAGAGATTGGAGCCCTTGAGCCCAT 239
 QY 1250 GGCAGCTCCAGCCTCTGCTCTCCAGAACTCAGCCCTCAGAGCTTAAGCAGCATGGA 1309
 Db 240 GGCAGCTCCAGCCTCTGCTCTCCAGAACTCAGCCCTCAGAGCTTAAGCAGCATGGA 299
 QY 1310 CCGGCCATGTGGAGCGCCTCTCAGCTTGAGCCGCTGCTTGCTCCAGGGGAGCCA 1369
 Db 300 CCGGCCATGTGGAGCGCCTCTCAGCTTGAGCCGCTGCTTGCTCCAGGGGAGCCA 359
 QY 1370 GGGGGCCCTCTGTTAGTACCCAAAGCGAGAGCGGATGCTTAATGAAGACAGTAGA 1429
 Db 360 GGGGGCCCTCTGTTAGTACCCAAAGCGAGAGCGGATGCTTAATGAAGACAGTAGA 419
 QY 1430 AGAGAAGACCTAGAGATTGAGAGCTTAAGACGAAGCAAAAGAACTGGAGGCCAAGAT 1489
 Db 420 AGAGAAGACCTAGAGATTGAGAGCTTAAGACGAAGCAAAAGAACTGGAGGCCAAGAT 479
 QY 1490 GTTGGCCCGAGAA 1501
 Db 480 GTTGGCCCGAGAA 491

RESULT 8

US-09-803-719-1937
 ; Sequence 1937, Application US/09803719
 ; Publication No. US2003004783A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Williams, Lewis T.
 ; APPLICANT: Escobedo, Jaime
 ; APPLICANT: Innis, Michael A.
 ; APPLICANT: Garcia, Pablo Dominquez
 ; APPLICANT: Sudduth-Klinger, Julie
 ; APPLICANT: Reinhard, Christoph
 ; APPLICANT: Giese, Klaus
 ; APPLICANT: Randazzo, Filippo
 ; APPLICANT: Kennedy, Giulia C.
 ; APPLICANT: Pot, David
 ; APPLICANT: Kassam, Altaf
 ; APPLICANT: Lamson, George
 ; APPLICANT: Drmanac, Radoje
 ; APPLICANT: Crkvenjakov, Radomir
 ; APPLICANT: Dickson, Mark
 ; APPLICANT: Drmanac, Snezana
 ; APPLICANT: Labat, Ivan
 ; APPLICANT: Leshkowitz, Dena
 ; APPLICANT: Kita, David
 ; APPLICANT: Garcia, Veronica
 ; APPLICANT: Jones, Lee William
 ; APPLICANT: Stache-Crain, Birgit
 ; TITLE OF INVENTION: Human Genes and Gene Products
 ; FILE REFERENCE: 1624.002
 ; CURRENT APPLICATION NUMBER: US/09/803,719
 ; CURRENT FILING DATE: 2001-03-09
 ; PRIOR APPLICATION NUMBER: 60/188,609
 ; PRIOR FILING DATE: 2000-03-09
 ; NUMBER OF SEQ ID NOS: 2396
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1937
 ; LENGTH: 386
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-803-719-1937

Query Match 24.5%; Score 376.4; DB 10; Length 386;
 Best Local Similarity 98.4%; Pred. No. 1.1e-104;

Matches 380; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1147 CTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTCCTTGGTCCACAGAGCAAGAGAG 1206
 Db 1 CGTTCTGCTGTTAAGCTGTCTCAGAAAGAAATTCCTTGGTCCACAGAGCAAGAGAGAG 60
 QY 1207 CCGAGGCCCCTGAGGAAGAGAGATTGGAGCCCTTGAGCCCATGAGCAGCTCCAGCCTCTG 1266
 Db 61 CCGAGGCCCCTGAGGAAGAGAGATTGGAGCCCTTGAGCCCATGAGCAGCTCCAGCCTCTG 120
 QY 1267 CCTCCAGAAACTCAGCCCTCAGAGAGCTTAAGCAGCATGAGCCGCGCATGCTGGAGC 1326
 Db 121 CCTCCAGAAACTCAGCCCTCAGAGAGCTTAAGCAGCATGAGCCGCGCATGCTGGAGC 180
 QY 1327 GCCTCCTCAGCTTGAGCCGCTGCTGCTCCAGGGGAGCAGGGGCCCTCTGTTGA 1386
 Db 181 GCCTCCTCAGCTTGAGCCGCTGCTGCTCCAGGGGAGCAGGGGCCCTCTGTTGA 240
 QY 1387 GTACCCCAAGCGAGAGCGGATGCTTAATGAAGACAGTAGAAGAGAGACCTTAGAGA 1446
 Db 241 GTACCCCAAGCGAGAGCGGATGCTTAATGAAGACAGTAGAAGAGAGACCTTAGAGA 300
 QY 1447 TTGAGAGCTTAAGCAAGCAAAAGAACTGGAGGCCCAAGATGTTGCCCAAGAGGCTG 1506
 Db 301 TTGAGAGCTTAAGCAAGCAAAAGAACTGGAGGCCCAAGATGTTGCCCAAGAGGCTG 360
 QY 1507 AGGAAGAGGAGAACCATTTGCCCCA 1532
 Db 361 AGGAAGAGGAGAACCATTTGCCCCA 386

RESULT 9

US-10-087-192-1180
 ; Sequence 1180, Application US/10087192
 ; Publication No. US20020182586A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Engelhard, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 ; TITLE OF INVENTION: CANCER
 ; FILE REFERENCE: 529452000122
 ; CURRENT APPLICATION NUMBER: US/10/087,192
 ; CURRENT FILING DATE: 2002-03-01
 ; PRIOR APPLICATION NUMBER: US 09/747,377
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; PRIOR FILING DATE: 2001-03-02
 ; NUMBER OF SEQ ID NOS: 2059
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1180
 ; LENGTH: 24525
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-087-192-1180

Query Match 23.7%; Score 365; DB 13; Length 24525;
 Best Local Similarity 73.8%; Pred. No. 1.4e-100;
 Matches 597; Conservative 0; Mismatches 0; Indels 212; Gaps 2;
 QY 550 AGGTATTAGACCTCCTCGAGCCCTGCTTCGGGAGACCTGGTAATCCGAGAGACTGCCGGG 609
 Db 2313 AGGTATTAGACCTCCTCGAGCCCTGCTTCGGGAGACCTGGTAATCCGAGAGACTGCCGGG 2372
 QY 610 GGAATATCCTGATTCCTGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTGAGC 669
 Db 2373 GGAATATCCTGATTCCTGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTGAGC 2432
 QY 670 GGCACTTCTGCGCAGCAGTCGAATCGGACTGTAGAGCCACCGGCTCAACAGCGCT 729
 Db 2433 GGCACTTCTGCGCAGCAGTCGAATCGGACTGTAGAGCCACCGGCTCAACAGCGCT 2492
 QY 730 CCTCCCGCAGTCATGCTGTGCTCCTGCTCA----- 759

Db 2493 CCTCCGACATCATGTGTCTCTCTGCTCAAGGTGAGGCGCAGACAGAGGGCGGAGGACCT 2552
 QY 760 -----AGGTGGACCA 770
 Db 2553 GGGAGCCAGGAGCCTGAGCTAAGCAGAGACCTTTGTTCTTACCCCGAGGTGGACCA 2612
 QY 771 CGGGAACGTTTGGGCCCATTTCCGACGAGAGGGAACCTCTACCTGATTGACTTGGCT 830
 Db 2613 CGGGAACGTTTGGGCCCATTTCCGACGAGAGGGAACCTCTACCTGATTGACTTGGCT 2672
 QY 831 GGTCAGAGAGCAACCGGCGCAGAGCAAGGCGCTTCGGCTAAAGAGAGTGAAGCC 890
 Db 2673 GGTCAGAGAGCAACCGGCGCAGAGCAAGGCGCTTCGGCTAAAGAGAGTGAAGCC 2732
 QY 891 ATCAACACCTCCCTGTTGTTGCTTGGGCAAGTGGTAGTGGCTGAATCAGGGCGCTCCCT 950
 Db 2733 ATCAACACCTCCCTGTTGTTGCTTGGGCAAGTGGTAGTGGCTGAATCAGGGCGCTCCCT 2792
 QY 951 CGTGTACCTTATCGGACAGCAAGCTCACCTCGCTATTG----- 989
 Db 2793 CGTGTACCTTATCGGACAGCAAGCTCACCTCGCTATTGCGCTATTGCGCTCAGGCCACCTGTCTC 2852
 QY 990 ----- 999
 Db 2853 AGGGAAGAGGGGCTGCAGAAAGAGGTTCTCAGGCGCTGCTGTGGGTGGGGAATAGCAGT 2912
 QY 990 -----CAGGACTC 997
 Db 2913 TGAGGCATAGAAAGGTGGGCTTCTGACCCACCACTGCGTGTCTCACTCAGGACTC 2972
 QY 998 TCTGGGTGGCTCAGCCACAGTATCTTATTGCAACATTTGCCATTTGCCCTGAGAGAGCTTCTA 1057
 Db 2973 TCTGGGTGGCTCAGCCACAGTATCTTATTGCAACATTTGCCATTTGCCCTGAGAGAGCTTCTA 3032
 QY 1058 CTAGACACAGTCTCCGCACTCACTTTGCTGCCAGGTCAAGAGGTGATCAATCGGCC 1117
 Db 3033 CTAGACACAGTCTCCGCACTCACTTTGCTGCCAGGTCAAGAGGTGATCAATCGGCC 3092
 QY 1118 TTTTACCAATGAGAGCTCGAGCTCATG 1146
 Db 3093 TTTTACCAATGAGAGCTCGAGCTCATG 3121

RESULT 10

US-09-918-995-31705
 ; Sequence 31705, Application US/09918995
 ; Publication No. US20030073623A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc.
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
 ; FROM VARIOUS CDNA LIBRARIES
 ; FILE REFERENCE: 20411-756
 ; CURRENT APPLICATION NUMBER: US/09/918,995
 ; PRIOR FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: US/09/235,076
 ; PRIOR FILING DATE: 1999-01-20
 ; NUMBER OF SEQ ID NOS: 38054
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 31705
 ; LENGTH: 464
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(464)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-918-995-31705

Query Match 17.2%; Score 265; DB 10; Length 464;
 Best Local Similarity 77.8%; Pred. No. 1.7e-70;
 Matches 333; Conservative 0; Mismatches 95; Indels 1; Gaps 1;
 QY 200 GCGGGCATGGACAGCTGTCTCTAGAGATTGCTAACTGGAGGAACCAACGAGGAGACTCT 259

Db 37 GCCTGAATTCATATATCGCGGTGATAGAGTGTAAACGCGCCGAAACCAACCCGATAATCT 96
 QY 260 CAATACCAAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGAGG 319
 Db 97 CCCTTACCTGTGCTTTGCCCTTTGACAGCCATAGGAGGACTGCGCTCTGACATGTACCTTCC 156
 QY 320 TTCAGTGCAGCCCATCTTAAGGCACCTTGTGGAAGGCGCAAGTGCAGAGTGTGCTTGCCTA 379
 Db 157 TTGA CCGTGGACCGTGCCACTGCTCTTTTCTGTAATAAAGCAGGAGGAGCTGGCCCTA 216
 QY 380 TGGACCCACAGGAGCTGGGAAGCAGCACAAATGCTGGGCGAGCCAGAGCAACCTGGGGT 439
 Db 217 CGGTGCTGAGGAGCTGGCATTAC-ATCTCGATGCTTGGCAGCCAGAGCAACCTGGGGT 275
 QY 440 GATCCCGCGGGCTCTCATGGACCTCTCAGAGCTCAAGGGAGGAGGGTGCAGAGGCCG 499
 Db 276 GATCCCGCGGGCTCTCATGGACCTCTCAGAGCTCAAGGGAGGAGGGTGCAGAGGCCG 335
 QY 500 GCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTAATCTCT 619
 Db 336 GCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTAATCTCT 395
 QY 560 CTTCTGAGACCTGCTTCCGGAGACCTGCTTAATCCGAGAGACTGCGCGGGGAATATCTCT 619
 Db 396 CTTCTGAGACCTGCTTCCGGAGACCTTGGAAATCCGAGAGACTGCGCGGGGAATATCTCT 455
 QY 620 GATTCGGG 628
 Db 456 GATTCGGG 464

RESULT 11

US-10-087-192-1177
 ; Sequence 1177, Application US/10087192
 ; Publication No. US20020182586A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Engelhard, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 ; FILE REFERENCE: 529452000122
 ; CURRENT APPLICATION NUMBER: US/10/087,192
 ; CURRENT FILING DATE: 2002-03-01
 ; PRIOR APPLICATION NUMBER: US 09/747,377
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; PRIOR FILING DATE: 2001-03-02
 ; NUMBER OF SEQ ID NOS: 2059
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1177
 ; LENGTH: 24291
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(24291)
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-087-192-1177

Query Match 17.1%; Score 262.4; DB 13; Length 24291;
 Best Local Similarity 73.8%; Pred. No. 4e-69;
 Matches 385; Conservative 0; Mismatches 61; Indels 76; Gaps 1;
 QY 550 AGGTATTAGACCTTCTGGACCTGCTTCCGGAGAGCTGTAATCCGAGAGAGTGCCTGGG 609
 Db 2939 AGGTATTAGACCTTCTGGATCTGATCAGGAGACCTCTGATCCGGAAGACTGCGGAG 2998
 QY 610 GGAATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTCTTGCTGATTGAGC 669
 Db 2999 GGAACATCTGATCCAGGCCTCACACAGAGCCCATCACCAGCTTCTCTGACTTCGAGC 3058
 QY 670 GGCACCTTCTGCGCAGCAGTCCGAATCGGACTGTAGGAGCCACCGGCTCAACACGAGCT 729

Db 3059 AGCATTCTTCCAGCCAGTAGAATCGAGCTGTAGGAGCCACCGGCTTAACCGAGCGT 3118
Qy 730 CCTCCGCGAGTATGCTGTGCTCTCTGGTCA-----759
Db 3119 CTTCCGCTAGTACGAGTGTCTTGGTCAAGGTAAAGGCCCTGACAGAAAGGCTGGG 3178
Qy 760 -----AGTGGACAGCGG 773
Db 3179 AAGCCCCGAAATGTGGACTGAGTCTGGTCTCTTCTTCCCTAGTAGATCAGCGT 3238
Qy 774 GAACCTTTGGCCCATTTCCGAGGAGGAGGAAACTCTACCTGATTCAGCTTTGGCTGGG 833
Db 3239 GAACCTTTGACTCCATTTCCGAGGAGGAGGAAACTCTACCTGATTCAGCTTTGGCTGGT 3298
Qy 834 TCAGAGGACAAACCGCGCAGAGGCAACAAAGGSCCTTCGCTAAAGAGAGTGGAGCCATC 893
Db 3299 TCAGAGGACAAACCGTCCGACAGCAACGAGCATTCGGCTCAAGAGAGCGGAGCCATC 3358
Qy 894 AACACCTCCCTGTTTGTCTGGGCAAGTGTAGATGCGCTGAATCAGGCGCTCCCTCGT 953
Db 3359 AACACCTCCCTGTTGTACTGGCAAGTGTGGATGATTAACAGGCGCTCCCTCGC 3418
Qy 954 GTACCTTATCGGACAGCAAGCTCATCTGCTTATTCAGGAC 995
Db 3419 ATACCATACCGGACAGCAAGCTCATCTGCTGTCAGGTC 3460

RESULT 12
US-10-029-386-10603
; Sequence 10603, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 10603
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR16.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
; OTHER INFORMATION: NT HIT: AB017334.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q14807, EVALUE 2.00e-36
; OTHER INFORMATION: EST_HUMAN HIT: BF975048.1, EVALUE 1.00e-126
US-10-029-386-10603

Query Match 15.2%; Score 233.2; DB 15; Length 531;
Best Local Similarity 98.7%; Pred. No. 9.9e-61;
Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 758 CAAGTGACACAGCGGGAACGTTTGGCCCCCATTTCCGAGGAGGAGGAAACTCTACCT 817
Db 34 CAAGTGACACAGCGGGAACGTTTGGCCCCCATTTCCGAGGAGGAGGAAACTCTACCT 93
Qy 818 GATTGACTTGGCTGGGTTCAGAGGACAAACCGGCGCACAGGCAACAGGCGCTTCGGCTAAA 877
Db 94 GATTGACTTGGCTGGGTTCAGAGGACAAACCGGCGCACAGGCAACAGGCGCTTCGGCTAAA 153
Qy 878 AGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTTGGGCAAGTGGTAGATCGGCTGAA 937
Db 154 AGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTTGGGCAAGTGGTAGATCGGCTGAA 213

Qy 938 TCAGGGCTCCCTCGTGTACCTTATCGGACAGCAAGCTCACTCGCCTATTTCAGGAC 995
Db 214 TCAGGGCTCCCTCGTGTACCTTATCGGACAGCAAGCTCACTCGCCTATTTCAGGAC 271
RESULT 13
US-10-029-386-24303
; Sequence 24303, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24303
; LENGTH: 232
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR16.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
; OTHER INFORMATION: SWISSPROT HIT: Q14807, EVALUE 5.00e-37
; OTHER INFORMATION: EST_HUMAN HIT: BE382882.1, EVALUE 1.00e-125
; OTHER INFORMATION: NT HIT: gi6453817, EVALUE 1.00e-125
US-10-029-386-24303

Query Match 15.0%; Score 230.4; DB 15; Length 232;
Best Local Similarity 99.6%; Pred. No. 5.4e-60;
Matches 231; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 762 GTGGACACAGCGGGAACGTTTGGCCCCCATTTCCGAGGAGGAGGAAACTCTACCTGATT 821
Db 1 GTGGACACAGCGGGAACGTTTGGCCCCCATTTCCGAGGAGGAGGAAACTCTACCTGATT 60
Qy 822 GACTTGGCTGGGTTCAGAGGACAAACCGGCGCACAGGCAACAGGCGCTTCGGCTAAAAGAG 881
Db 61 GACTTGGCTGGGTTCAGAGGACAAACCGGCGCACAGGCAACAGGCGCTTCGGCTAAAAGAG 120
Qy 882 AGTGAGGACCATCAACACCTCCCTGTTTGTCTTGGGCAAGTGGTAGATGCGCTGAATCAG 941
Db 121 AGTGAGGACCATCAACACCTCCCTGTTTGTCTTGGGCAAGTGGTAGATGCGCTGAATCAG 180
Qy 942 GGCCTCCCTCGTGTACCTTATCGGACAGCAAGCTCACTCGCCTATTTCGAGG 993
Db 181 GGCCTCCCTCGTGTACCTTATCGGACAGCAAGCTCACTCGCCTATTTCGAGG 232

RESULT 14
US-10-108-260A-249
; Sequence 249, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1e1 full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 249
; LENGTH: 2095
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-249


```
QY 395 TGGGAAGACGACACAAATGCTGGGACGCCAGAGCAACCTGGGGTGATCCCGGGGCTCT 454
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 612 TGGGAACCTACACCATGCTGGGCACAGACCAGGAGCTGGCACTATGTTTCAGACCCCT 671
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 455 CATGACCTCTCTGAGCTCACAAGGAGAGAGGTGCGAGGGCCGGCCATGGGGCCCTTTC 514
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 515 TGTCAACATGCTTACCTAGAGATCTACCAGAGAGAGGTATTAGACCTCTCGAGCCCTGC 574
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 723 GGTCTCCATGCTCCTCTGAGATCTAATGAGATGATCCGGGACCTGCTGAACCCCTC 782
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 575 TTCGGGAGACCTGGTAATCCGAGAAGACTCGCGGGGAATATCCTGATTCGGGGTCTCTC 634
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 783 CCTGGCTACCTGGAGCTCGCGGAGGACTCTAAGGGGTGATCCAGGTGGCCGGCATCAC 842
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 635 CCAGAAGCCCATCAGTAGCTTTGCTGATTTTGAAGGGCACTTCTGCGAGCCAGTCGAAA 694
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 843 CGAAGTCTCCACCATCAATGCCAAGGAGATCATGAGCTGTCTGATGAAGGGGAACCGCA 902
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 695 TCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCCCGCAGTCTGCTGTCTCT 754
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 903 GAGGACCCAGGAGCCACCGCGCCCAACAGAGCTCTCCCGCTCCACGCGGTACTGCA 962
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 755 GGTCAAGGTGACACAGCGGGAACGTTT---GGCCCCATTTCCCGCAGAGAGGGGAACCT 811
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 963 GGTGACCGTGGCCAGCGAGCGCGGTCAAGAACATCTTGAGGAGCGCGCGGCT 1022
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 812 CTACCTGATGATCTTGGCTGGGTGAGAGCAACCGGGGCAAGGCAACAAAGGGCTTCG 871
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1023 GTTCATGATCGACCTGGCTGAGAGCGCGCTCGCAGACACAGAAATCGTGGGCGCG 1082
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 872 GCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTCTCTGGGCAAGTGGTAGATGC 931
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1083 TATGAAGAGGGGGCCCATCAACCGCTCACTGTGGCACTGGGCAACTGCATCAACGC 1142
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 932 GCTGAATCAGG-----GCTCCCTCGTGTACCTTATCGGACAGCAAGCTCAGCT 985
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1143 CCTGAGCGACAGGGTAGCAACAGTACATCACTATCGGACAGCAAGCTCAGCGGCT 1202
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 986 ATTGAGGACTCTCTGGGTGGCTCAGCCCAAGTATCCTTATGCAACATTGCCCTTGA 1045
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1203 CCTGAAGGACTCTCTGGGAGGAAACAGCGGACAGTATGATCGCTCACATCAGTCTTGC 1262
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1046 GAGACGCTTCTACCTAGACAGTCTCGCACTCACTTTGCTGCCAGTCCAGGAGGT 1105
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1263 GAGCAGTGCCTTCGAGGAGTCCGGAAACACCTGACCTACGCCGCGCGCCAGAACAT 1322
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1106 GATCAATCGG 1115
Db ||||| |||||
QY 1323 TAAGACTAGG 1332
Db ||||| |||||
```

Search completed: November 11, 2004, 01:20:38
Job time : 825.883 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 04:25:02 ; Search time 5346.62 Seconds
(without alignments)
10482.189 Million cell updates/sec

Title: US-10-797-893-3

Perfect score: 1538

Sequence: 1 atgcagcgcggcgctcg.....accattgtccacaatgtga 1538

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1.*

2: gb_est2.*

3: gb_hic.*

4: gb_est3.*

5: gb_est4.*

6: gb_est5.*

7: gb_est6.*

8: gb_gse1.*

9: gb_gse2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1526.8	99.3	2057	3	CR615781 full-leng
2	1512.8	98.4	2025	3	CR590974 full-leng
3	1462.8	95.1	2016	3	CR590954 full-leng
4	1074.8	69.9	2055	3	AK084310 Mus muscu
5	1070.6	69.6	2056	3	AK075922 Mus muscu
6	1054.4	68.6	1124	1	AL545712 AL545712
7	1002.8	65.2	1027	5	BX337406 BX337406
8	980.8	63.8	1083	1	AL559772 AL559772
9	978.6	63.6	1032	5	BX337353 BX337353
10	942	61.2	959	5	BX363358 BX363358
11	934	60.7	958	5	BX354166 BX354166
12	922.4	60.0	1031	4	BM910587 AGENCOURT
13	921.4	59.9	999	1	AL559769 AL559769
14	921.4	59.9	1094	5	BX376199 BX376199
15	909	59.1	1019	4	BM546877 AGENCOURT
16	907.8	59.0	978	5	BX400806 BX400806
17	905	58.8	1016	1	AL518361 AL518361
18	899.8	57.9	1051	4	BM806996 AGENCOURT
19	899.6	57.8	980	5	BX404364 BX404364
20	889.2	57.8	1041	4	BM478215 AGENCOURT
21	887.4	57.7	915	5	BX383181 BX383181
22	882.8	57.4	1034	5	BQ051730 AGENCOURT
23	875.4	56.9	971	5	BQ053671 AGENCOURT
24	862.4	56.1	926	1	AL578338 AL578338

25	861.8	56.0	1040.	4	BM546897 AGENCOURT
26	861.6	56.0	991	5	BX333336 BX333336
27	851.8	55.4	865	1	AL555550 AL555550
28	851.4	55.4	865	6	CD109092 AGENCOURT
29	844	54.9	990	5	BX399192 BX399192
30	836.4	54.4	1139	4	BM553107 AGENCOURT
31	833	54.2	838	5	BX341048 BX341048
32	831.8	54.1	851	5	BUI188590 AGENCOURT
33	831.6	54.1	942	5	BUS27647 AGENCOURT
34	823	53.5	995	1	AL559752 AL559752
35	821.8	53.4	846	5	BX404325 BX404325
36	821.4	53.4	1640	3	AK088033 Mus muscu
37	820.8	53.4	978	4	BI754589 603025126
38	819.4	53.3	981	4	BM561588 AGENCOURT
39	819	53.3	921	5	BUS60011 AGENCOURT
40	816.4	53.1	878	1	AL576578 AL576578
41	815.2	53.0	872	5	BUS21359 AGENCOURT
42	813	52.9	908	5	BQ953445 AGENCOURT
43	808.6	52.6	873	5	BQ955064 AGENCOURT
44	808.6	52.6	971	5	BX397059 BX397059
45	802.6	52.2	845	4	BG686012 602638562

ALIGNMENTS

RESULT 1	CR615781	2057 bp	mrna	linear	HTC 21-JUL-2004
LOCUS	full-length cDNA clone CSODK012YE23 of HeLa cells Cot 25-normalized				
DEFINITION	of Homo sapiens (human).				
ACCESSION	CR615781				
VERSION	CR615781.1	GI:50496588			
KEYWORDS	HTC; CNSLT_cDNA.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	Li, W.B., Gruber, C., Jesse, J., and Polayes, D.				
JOURNAL	Full-length cDNA libraries and normalization				
REMARK	Unpublished				
	Contact : Feng Liang Email : fliang@lifetech.com URL :				
	http://fulllength.invitrogen.com/ Invitrogen Corporation 1600				
REFERENCE	Paraday Avenue				
AUTHORS	2 (bases 1 to 2057)				
TITLE	Genoscope.				
JOURNAL	Direct Submission				
	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :				
	BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr				
	- Web : www.genoscope.cns.fr)				
COMMENT	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime				
	end enriched, double-strand cDNA was digested with Not I and cloned				
	into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library				
	was normalized. Library was constructed by Life Technologies, a				
	division of Invitrogen.				
FEATURES	Location/Qualifiers				
source	1..2057				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="CSODK012YE23"				
	/tissue_type="HeLa cells Cot 25-normalized"				
	/plasmid="pCMVSPORT_6"				

ORIGIN	Query Match	99.3%	Score 1526.8;	DB 3;	Length 2057;
	Best Local Similarity	99.9%	Pred. No. 0;		
	Matches 1528;	Conservative	0;	Mismatches	2;
				Indels	0;
				Gaps	0;
Qy	6	GCCCGCGCGCGCTCGACGCGAGCAGCGAGCGAGATGGCGGAGCTTACGCGCGCG	65		
Db	7	GCCCGCGCGCGCTCGACGCGAGCAGCGAGCGAGATGGCGGAGCTTACGCGCGCG	66		


```
QY 20 GACGAGCAGAGCGACCGAGATGGCGGAGCTTTCAGCGCGCGATCTTCAGAGCTGG 79
Db 1 GACGAGCAGAGCGACCGAGATGGCGGAGCTTTCAGCGCGCGATCTTCAGAGCTGG 60
QY 80 TCGCTGTGGCTTAAGCAAGATTGGAGCTACTCTGCTGCTTCCACTTCAGCTCCGCTAAGGCT 139
Db 61 TCGCTGTGGCTTAAGCAAGATTGGAGCTACTCTGCTGCTTCCACTTCAGCTCCGCTAAGGCT 120
QY 140 GGCTGTGGCACTGGCGCCATTTTGGATGGAAACAGCGGGAGCAAGTGAATCCCCCTGTGT 199
Db 121 GGCTGTGGCACTGGCGCCATTTTGGATGGAAACAGCGGGAGCAAGTGAATCCCCCTGTGT 180
QY 200 CGCGGGCATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACTCT 259
Db 181 CGCGGGCATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACTCT 240
QY 260 CAATATCCAGTTTGAATGCTTCTATGGGGAGAGAGTACTCAGCGAGACTCTATGCGAG 319
Db 241 CAATATCCAGTTTGAATGCTTCTATGGGGAGAGAGTACTCAGCGAGACTCTATGCGAG 300
QY 320 TTCACTGAGCCCATCTTAAGGCACTTCTGGAAGGGCAGAAATGCCAGTGTGCTTGCTTA 379
Db 301 TTCACTGAGCCCATCTTAAGGCACTTCTGGAAGGGCAGAAATGCCAGTGTGCTTGCTTA 360
QY 380 TGGACCCACAGAGCTGGGAAGACGACACAAATGCTGGGCGAGCCAGAGCAACCTGGGGT 439
Db 361 TGGACCCACAGAGCTGGGAAGACGACACAAATGCTGGGCGAGCCAGAGCAACCTGGGGT 420
QY 440 GATCCGCGGGCTCTCATGGAACCTCTGCACTCTCAAGGGAGAGGGTCCGAGGGCG 499
Db 421 GATCCGCGGGCTCTCATGGAACCTCTGCACTCTCAAGGGAGAGGGTCCGAGGGCG 480
QY 500 GCCATGGGCCCTTCTGTACCATGCTTACTAGAGATCTACCAGAGAGAGTATTAGA 559
Db 481 GCCATGGGCCCTTCTGTACCATGCTTACTAGAGATCTACCAGAGAGAGTATTAGA 540
QY 560 CCTCTGAGACCTGCTTGGGAGACCTGGTAAATCCGAGAGACTGCCGGGGGAATATCCT 619
Db 541 CCTCTGAGACCTGCTTGGGAGACCTGGTAAATCCGAGAGACTGCCGGGGGAATATCCT 600
QY 620 GATTCGGGCTCTCCAGAGAGCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTCCT 679
Db 601 GATTCGGGCTCTCCAGAGAGCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTCCT 660
QY 680 GCCAGCAGTCGAAATCGGACTGTAGAGCACCCTGCTCAACAGGCTCTCTCCCGAG 739
Db 661 GCCAGCAGTCGAAATCGGACTGTAGAGCACCCTGCTCAACAGGCTCTCTCCCGAG 720
QY 740 TCATGCTGTGCTCTGCTCAAGGTGGACACAGCGGGAACGTTTGGCCCCCATTTGCCAGCG 799
Db 721 TCATGCTGTGCTCTGCTCAAGGTGGACACAGCGGGAACGTTTGGCCCCCATTTGCCAGCG 780
QY 800 AGAGGGAATCTTACTGATGATGCTTGGCTGGGTTCAGAGCAACCGGGCGCAGGCAA 859
Db 781 AGAGGGAATCTTACTGATGATGCTTGGCTGGGTTCAGAGCAACCGGGCGCAGGCAA 840
QY 860 CAAGGGCTTCGGCTTAAAGAGATGGAGCCATCAACACCTCCCTGTTGCTCGGCAA 919
Db 841 CAAGGGCTTCGGCTTAAAGAGATGGAGCCATCAACACCTCCCTGTTGCTCGGCAA 900
QY 920 AGTGGTAGTCGCTGAATCAGGGCCCTCCCTGCTGTATCTTATCGGACAGCAAGCTCAC 979
Db 901 AGTGGTAGTCGCTGAATCAGGGCCCTCCCTGCTGTATCTTATCGGACAGCAAGCTCAC 960
QY 980 TCGCTTATGAGGACTCTCTGGGTGGCTCAGCCACAGTATCTTATGCGCAACATGC 1039
Db 961 TCGCTTATGAGGACTCTCTGGGTGGCTCAGCCACAGTATCTTATGCGCAACATGC 1020
QY 1040 CCCTGAGAGAGCTTCTACCTAGACAGAGTCTCGGCACTCACTTCTGCGAGGCTCAA 1099
Db 1021 CCCTGAGAGAGCTTCTACCTAGACAGAGTCTCGGCACTCACTTCTGCGAGGCTCAA 1080
QY 1100 GGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTCAGGCTCTATGCTTGGGACCTGT 1159
```

```
Db 1081 GGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTCAGGCTCTATGCTTGGGACCTGT 1140
QY 1160 TAAGCTGTCTCAGAAAGAAATTTGCTTGTGCTCCACAGAGCAAGAGAGCCCGAGCCCTGA 1219
Db 1141 TAAGCTGTCTCAGAAAGAAATTTGCTTGTGCTCCACAGAGCAAGAGAGCCCGAGCCCTGA 1200
QY 1220 GGAAGAGAGATTTGGGAGCCCTGAGCCCATGGCAGCTTCAGCCCTTGCCTCCAGAAACT 1279
Db 1201 GGAAGAGAGATTTGGGAGCCCTGAGCCCATGGCAGCTTCAGCCCTTGCCTCCAGAAACT 1260
QY 1280 CAGCCCCCTACAGAAAGCTAAGCAGCATGGACCCCGCCATGCTGGAGCGCTCTCTCAGCTT 1339
Db 1261 CAGCCCCCTACAGAAAGCTAAGCAGCATGGACCCCGCCATGCTGGAGCGCTCTCTCAGCTT 1320
QY 1340 GGACCGTCTGCTTGCCTCCAGGGGAGCAGGGGGCCCTCTGTTAGTAGTACCCCAAGCG 1399
Db 1321 GGACCGTCTGCTTGCCTCCAGGGGAGCAGGGGGCCCTCTGTTAGTAGTACCCCAAGCG 1380
QY 1400 AGACCGATGCTGCTTAATGAAGACAGTAGAAGAGAGAGACCTAGAGATTGAGAGGCTTAA 1459
Db 1381 AGACCGATGCTGCTTAATGAAGACAGTAGAAGAGAGAGACCTAGAGATTGAGAGGCTTAA 1440
QY 1460 GACGAACAAAGAACTGGAGGCCAAGATGTTGCCCCAGAAAGCTGAGGAAAAGGAGAA 1519
Db 1441 GACGAACAAAGAACTGGAGGCCAAGATGTTGCCCCAGAAAGCTGAGGAAAAGGAGAA 1500
QY 1520 CCATTGCTCCCAATG 1535
Db 1501 CCATTGCTCCCAATG 1516

RESULT 3
CR590954
LOCUS CR590954 2016 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0D1040V106 of Placentia Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR590954
VERSION CR590954.1 GI:50471761
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2016)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 2016)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES Location/Qualifiers
source
1. .2016
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1040V106"
/tissue_type="Placentia Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 95.1%; Score 1462.8; DB 3; Length 2016;
```

Best Local Similarity	99.9%;	Pred. No. 0;	
Matches 1464;	Conservative	0;	Mismatches
		2;	Indels
		0;	Gaps
		0;	Gaps

QY	70	CAGGAGCTGGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCTGTCGTCCACCTCCAGCTC	129
Db	1	CAGGAGCTGGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCTGTCGTCCACCTCCAGCTC	60
QY	130	GCGTAAAGGGTGGCTGTCGCACTCGCGGCATTTTGGATGGAACAGCGGGAGCAAGTCATC	189
Db	61	GCGTAAAGGGTGGCTGTCGCACTCGCGGCATTTTGGATGGAACAGCGGGAGCAAGTCATC	120
QY	190	CCCCCTGTGTGCGGGGATGCACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACC	249
Db	121	CCCCCTGTGTGCGGGGATGCACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACC	180
QY	250	AGGAGACTCCTAAATACACAGTTTGATGCCCTTCTATGGGGAGAGAGTACTCAGCAGGACA	309
Db	181	AGGAGACTCCTAAATACACAGTTTGATGCCCTTCTATGGGGAGAGAGTACTCAGCAGGACA	240
QY	310	TCTATGACAGTTTCAGTGCAGCCCATCTTAAGGCACTTGTCTGGAAGGCGAGAATGCCAGTG	369
Db	241	TCTATGACAGTTTCAGTGCAGCCCATCTTAAGGCACTTGTCTGGAAGGCGAGAATGCCAGTG	300
QY	370	TGCTTTGCTTATGGACCCACAGGAGCTGGGAAGACGCACACAATGCTGGGAGGCCACAGC	429
Db	301	TGCTTTGCTTATGGACCCACAGGAGCTGGGAAGACGCACACAATGCTGGGAGGCCACAGC	360
QY	430	AACCTGGGGTGATCCCGGGGCTCTCATGACCTCTCTGCACTCTCAAGGGAGAGGGGTG	489
Db	361	AACCTGGGGTGATCCCGGGGCTCTCATGACCTCTCTGCACTCTCAAGGGAGAGGGGTG	420
QY	490	CCGAGGSCCGGCCATATGGGCCCTTCTCTGACCATGCTTACCTAGAGATCTACACAGGAGA	549
Db	421	CCGAGGSCCGGCCATATGGGCCCTTCTCTGACCATGCTTACCTAGAGATCTACACAGGAGA	480
QY	550	AGGTATTAGACCTCTCTGGACCCCTGCTTCGGGGAGACCTGGTAATCCGAGAAGACTGCCGGG	609
Db	481	AGGTATTAGACCTCTCTGGACCCCTGCTTCGGGGAGACCTGGTAATCCGAGAAGACTGCCGGG	540
QY	610	GGAAATATCTTGATTCGGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTGATTTTGAGC	669
Db	541	GGAAATATCTTGATTCGGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTGATTTTGAGC	600
QY	670	GGCACTTTCCTGCGAGCCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACACAGCGCT	729
Db	601	GGCACTTTCCTGCGAGCCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACACAGCGCT	660
QY	730	CCTCCCGCAGTCATGCTGTGCTCCTGTGTAAGGTGGACAGCGGGAACTGTTGGCCCCCAT	789
Db	661	CCTCCCGCAGTCATGCTGTGCTCCTGTGTAAGGTGGACAGCGGGAACTGTTGGCCCCCAT	720
QY	790	TTCCGCGAGCAGAGGGGAAAATCTACTGATTGACTTTGGCTGGGTCAAGAGACACCGCGC	849
Db	721	TTCCGCGAGCAGAGGGGAAAATCTACTGATTGACTTTGGCTGGGTCAAGAGACACCGCGC	780
QY	850	GCACAGCAACAAGGGCCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTG	909
Db	781	GCACAGCAACAAGGGCCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTG	840
QY	910	TCCTGGGCAAGTGGTAGATCGCTGAATCAGGGCCCTCCCTCGTGTACTTATCGGGGACA	969
Db	841	TCCTGGGCAAGTGGTAGATCGCTGAATCAGGGCCCTCCCTCGTGTACTTATCGGGGACA	900
QY	970	GCAAGCTCACTCGGCTATTGAGGACTCTCTGGTGGCTCAGGCCACAGTATCGCTATTG	1029
Db	901	GCAAGCTCACTCGGCTATTGAGGAGTCTCTGGTGGCTCAGGCCACAGTATCGCTATTG	960
QY	1030	CCAACATTGCCCTTGAGAGAGCTTCTACCTAGACACAGTCTCGGCACTCAACTTTGCTG	1089
Db	961	CCAACATTGCCCTTGAGAGAGCTTCTACCTAGACACAGTCTCGGCACTCAACTTTGCTG	1020
QY	1090	CCAGGTCCTAAGGAGGTGATCAATCGGCCCTTTTACCAATGAGAGCCTGCACCTCATGCTC	1149

Db	1021	CCAGGTCCTCAAGGAGGTGATCAATCGCGCTTTTACCAATGAGAGCCTGACGCCTCATGCCT	108
Qy	1150	TGGGACCTGTTAAGCTGTCTCAGAAAAGAAATTGCTTGGTCCACCAGAGGCAAGAGAGAGCCC	1209
Db	1081	TGGGACCTGTTAAGCTGTCTCAGAAAAGAAATTGCTTGGTCCACCAGAGGCAAGAGAGAGCCC	1140
Qy	1210	GAGGCCCTGAGGAAGAGGAGATTGGGAGCGCCTGAGGCCCATGGCAGTCCAGCCTCTGCCT	1269
Db	1141	GAGGCCCTGAGGAAGAGGAGATCGGAGCGCCTGAGGCCCATGGCAGTCCAGCCTCTGCCT	1200
Qy	1270	CCAGAAACTCAGCCCTTACAGAACTAAGCAGCATGAGCCCGGCATGCTCGAGAGGCC	1329
Db	1201	CCAGAAACTCAGCCCTTACAGAACTAAGCAGCATGAGCCCGGCATGCTCGAGAGGCC	1260
Qy	1330	TCCTCAGCTTGGACCGTCTGCTTGGCTCCCGAGGGGAGCAGGGGGCCCTCTCTGTGAGTA	1389
Db	1261	TCCTCAGCTTGGACCGTCTGCTTGGCTCCCGAGGGGAGCAGGGGGCCCTCTCTGTGAGTA	1320
Qy	1390	CCCCAAAGCGAGAGCGGATGGTGCTTAATGAAGACAGTGAAGAAGAGGACCTAGAGATTG	1449
Db	1321	CCCCAAAGCGAGAGCGGATGGTGCTTAATGAAGACAGTGAAGAAGAGGACCTAGAGATTG	1380
Qy	1450	AGAGGCTTAAGACGAGCAAGCAAAAAGAACTGGAGGCCAAGATGTTGGCCCGAGAGGCTGAGG	1509
Db	1381	AGAGGCTTAAGACGAGCAAGCAAAAAGAACTGGAGGCCAAGATGTTGGCCCGAGAGGCTGAGG	1440
Qy	1510	AAAAGGAGAACCAATTGTCCACAAATG	1535
Db	1441	AAAAGGAGAACCAATTGTCCACAAATG	1466
RESULT 4			
AK084310		2055 bp	linear
LOCUS	AK084310	2055 bp	linear
DEFINITION	Mus musculus 12 days embryo eyeball cDNA, RIKEN full-length		HTC 03-APR-2004
ACCESSION	AK084310		
VERSION	AK084310.1	GI:26101874	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus		
ORGANISM	Mus musculus (house mouse)		
REFERENCE			
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20493374		
PUBMED	11042159		
REFERENCE			
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., S. Sumi, N., Iehi, Y., Nakamura, S., Hazama, M., Nishie, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system - 384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE			
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the		

TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
Nature 409, 685-690 (2001)
REFERENCE
5
AUTHORS
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL
Nature 420, 563-573 (2002)
REFERENCE
6 (bases 1 to 2055)
AUTHORS
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashi, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imokani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kohji, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitho, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE
Direct Submission
JOURNAL
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>
URL: <http://fantom.gsc.riken.jp/>

FEATURES
source
1..2055
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:D230020105"
/db_xref="taxon:10090"
/clone="D230020105"
/tissue_type="eyeball"
/clone_libs="RIKEN full-length enriched mouse cDNA library"
/dev_stages="12 days embryo"
8..1991
/note="kinesin superfamily protein KIF22 putative"

ORIGIN
Query Match 69.9%; Score 1074.8; DB 3; Length 2055;
Best Local Similarity 83.6%; Pred. No. 1.8e-273;
Matches 1214; Conservative 0; Mismatches 238; Indels 0; Gaps 0;

Qy 59 GCGCGGATCTCAGGAGTGTGCTGCGCTGCGCTAAGCAAGATTGAGACTACTCGTGTCC 118
Db 49 GCGGTGCGGCCACCTCAGGACCTGGCGCTGTGTGAAGCAAGAGGGCTTGGCCGCGTCC 108

Qy 119 ACCTCCAGTCGCTGAGGGTGGCTGTGCGACTCGGCCATTCTTGATGGAACAGCGGG 178
Db 109 GCCTCTAGCTCGAGTAAGGGTGGCTGTGCGCGCTGCGCCCGTTATGATGAGAGACAGA 168

Qy 179 AGCAAGTGAATCCCGCTGTGTGCGGGGCATCGACAGCTGCTCTCTAGAGATTGCTAACTG 238
Db 169 AGCGAAGGAGTCCCGCTGTGTGCGAGCCATAGACAGCTGCTCTCTGAGTGGCTAACTG 228

Qy 239 GAGGAACCCAGGAGACTCTCAATACCAATGCTTTGCTTATGCGGGAGAGGAGTAC 298
Db 229 GAAGAAATACCGAGAGCGCTCAAATATCAGTTTGTGCTTCTATGCGGAGAGAGACAC 288

Qy 299 TCAGCAGGACATCTATGCAAGTTTCAGTGCAGCCCATCTTAAGCCACTTGTCTGGAAGGCA 358
Db 289 TCAGCAGGAAGTCTATGTAGTTTCAGTACAGCCCATCTTAAGCCACTTGTCTGGAAGGCA 348

Qy 359 GAATGCCAGTGTGCTTGTCTATGAGACCACAGGAGCTGGGAAGACGACACAATGTGGG 418
Db 349 GAATGCCAGTGTACTTGTCTATGAGCCCTACTTGGGCGAGGAAGACACACACAATGTGGG 408

Qy 419 CAGCCGAGAGCAACTGGGGTGTATCCCGGGGCTCTCATGGACCTCTCGACGCTCACAG 478
Db 409 CAGCCGAGAGCAACTGGAGTGATTCCTCGGGCTCTCATGGACCTCTCGACGCTCACAG 468

Qy 479 GGAGGAGGCTGCCGAGGCGCGCCCATGAGCCCTTCTGTCTCAACCTGTCTTACTAGAGAT 538
Db 469 GGAAGAGAGTCTGAGGAGCGGCCCATGGAGCGTTCTGTCTGTATGTCTTATTTAGAGAT 528

Qy 539 CTACAGAGAGAGGTATTTAGACCTCTCTGGAACCTCTGCTTGGGAGACCTGGTAAATCCGAGA 598
Db 529 CTACAGAGAGAGGTATTTAGACCTCTTGGATCTCTGATCAGGAGACCTCGTGTATCCGGA 588

Qy 599 AGACTGCGGGGGAATATCTGATTCGGGTCTCTCCAGAGGCCCATCAGTACGTAGCTTTCG 658
Db 589 AGACTGCGGGGGAATATCTGATTCGGGTCTCTCCAGAGGCCCATCAGTACGTCTTC 648

Qy 659 TGATTTTTCAGGCGGCACTTCTCTGCGAGCCAGTCGAAATCGGACTGTAGGAGCCACCGGCT 718
Db 649 TGACTTTCAGGCGGCACTTCTCTGCGAGCCAGTCGAAATCGGACTGTAGGAGCCACCGGCT 708

Qy 719 CAACAGCGCTCTCCGCGAGTCACTGTCTCTCTGCTCAAGTGGACCGGGAACG 778
Db 709 TAACAGCGCTCTTCCCGTAGTCACGCGAGTCTCTTGTGTCAGGTAGATCAGCGTGAACG 768

Qy 779 TTTGGCCCCATTTGCGCAGCGAGAGGGAATACTCTACTGATTGACTTTGGCTGGGTCTAGA 838
Db 769 TTTGACTCTCAATTTGCGCAGCGGGAAGAAAACTCTACCTTATGATTGGTGGTCTAGA 828

Qy 839 GGACAAACCGGCGCAGCAGCAACAAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACAC 898
Db 829 GGACAAACCGTGCAGCAGCAACAAGGGCATTTGGCTAAAGAGAGCGGAGCCATCAACAC 888

Qy 899 CTCCTCTTTGTCTCTGGGCAAAAGTGTAGATGCGCTGAATCAGGCGCTCTCTCTGTGTACC 958
Db 889 CTCCTCTTTGTCTCTGGGCAAAAGTGTGTAGTGCATTTAAACAGGCGCTCTCTCTGTGTACC 948

Qy 959 TTATCGGAGCAGCAAGCTCACTCGCCTATTGCGAGACTCTCTGGGTGCTCAGCCACAG 1018
Db 949 ATACCGGAGCAGCAAGCTCACTCGCTCTGCTGCGAGACTCTCTGGGAGGCTCAGCTCATAG 1008

Qy 1019 TATCTTTATTGCCAAATTTGCCCTGAGAGAGCTTTTACCTAGACAGTCTCCGCACT 1078
Db 1009 CATCTCTATTGCCAAATTTGCTCTGAAAGAGCTTTTACAGGATACATTTTCAAGCATTT 1068

Qy 1079 CAACCTTTGCTGCCAGGTCCAAGAGGTGATCAATCGGCTTTTACCAATGAGAGCTGCA 1138
Db 1069 GAACCTTCACTGCTAGGTCCAAGAGGTGATTAACCGGCTTTTACCAATGAGAGCTTACA 1128

Qy 1139 GCCTCATGCTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTTGCTTGTTCACACAGAGC 1198
Db 1129 GCCTCATGCTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTTGCTTGTTCAGAGGC 1188

Qy 1199 AAGAGAGCCGAGGCGCTTGAAGAGAGGATTTGGAGCGCTTGGAGCCATGGGAGCTCC 1258
Db 1189 AAGAGAGCCGAGGCGCTTGAAGAGAAATCACTGGGAGCGCTTGTAGTCTACAGAGCTCC 1248

Qy 1259 AGCTCTGCTCTCCAGAAAATCTCAGCCCTTACAGAGCTAAGCAGATGGACCGGCCAT 1318
Db 1249 AGCTCTGCTCTCCAGAAAATCTAAGCTCTTACAGAGCTAAGCAATATGGATCCAGCCAT 1308

Qy 1319 GCTGAGAGCGCTCTCTAGCTTGGACCGTCTGCTTGTCTCCAGGGAGCCAGGGGGCCCC 1378
Db 1309 GTTGGAGAACCTCTGAGTATGGAACGTTTGTGGTTCCTCCAGGGAAGCCAGGAGACCCC 1368

Qy	599	AGACTGCCGGGGAAATATCTGTATTCGGGTCTCTCCAGAGGCCCATCTAGTAGCTTTGC	658
Db	592	AGACTGCCGGGGAAATATCTGTATTCGGGTCTCTCCAGAGGCCCATCTAGTAGCTTTGC	651
Qy	659	TGATTTTGGCGGCACATCTCTCCAGCGAGTTCGAATTCGGATCTGTAGAGCCACCGGCT	718
Db	652	TGATTTTGGCGGCACATCTCTCCAGCGAGTTCGAATTCGGATCTGTAGAGCCACCGGCT	711
Qy	719	CAACACGGCTCTCTCCCGCAGTCATGTGTCTCTCTGCTCAAGGTGCACAGCGGGAAACG	778
Db	712	TAACACGGCTCTCTCCCGTAGTCAGCAGTGTCTTTGGTCAAGGTAGATCAGCGTGAACG	771
Qy	779	TTTGGCCCCATTTCCGCGAGGAGGAAATCTCTACTCTGATTGACTTTGGCTGGGTGAGA	838
Db	772	TTTGTACTCCATTTCCGCGAGGAGGAAATCTCTACTCTGATTGATTGGCTGGTTCAGA	831
Qy	839	GGACAAACGGCGGACAGGCAACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACAC	898
Db	832	GGACAAACGGCGGACAGGCAACAGGGCTTCGGCTCAAGAGAGCGAGCCATCAACAC	891
Qy	899	CTCCCTCTTTGTCTCTGGCGAAAGTGTAGATGCCTGAATCAGAGGCTCTCCCTCGTGATCC	958
Db	892	CTCCCTCTTTGTACTGGGCANAGTGTGTGATGCATTAACCCAGGGCTCTCCCTCGCATACC	951
Qy	959	TTATCGGGACAGCAAGCTCACTCGCCTATTTCAGGACTCTCTGGGTGGCTCAGGCCACAG	1018
Db	952	ATACCGGGACAGCAAGCTCACTCGCTCTGCTGAGGACTCTCTGGGAGGCTCAGCTCATAG	1011
Qy	1019	TATCTTATTTCGCAACATTCGCCCTGAGAGACGCTTTACTAGACACAGTCTCCGCACT	1078
Db	1012	CATCTCTATTTCGCAACATTCGCTCTGAGAGACGCTTTTACAGGATCAATTTTCAGCATTT	1071
Qy	1079	CAACTTTGCTGCCAGGTCGCAAGGAGTGATCAATCGGCTTTTACCAATGAGAGCTGTACA	1138
Db	1072	GAACCTTCACTGCTAGGTCCAAGGAGGTGATTAACCGGCTTTTACCACCAATGAGAGCTGTACA	1131
Qy	1139	GCCTCATGCTCTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTCGTGGTCCACAGAGGC	1198
Db	1132	GCCTCATGCTCTGGACCTGTTAAGCTGTCTCAGAAAGAAATTCGTGGTCCACAGAGGC	1191
Qy	1199	AAAGAGAGCCGAGGCCCTCAGGAAAGAGAGATTGGAGCCCTGAGGCCCATGGCAGCTCC	1258
Db	1192	AAAGAGGCCAAAGGCCCTGAAGAAGAAATCAACTGGGAGCCCTGAGTCTACAGCAGCTCC	1251
Qy	1259	AGCCTCTGCTCTCCAGAAATCTACGCCCTTACAGAACTTAAGCAGCATGAGACCGGCCAT	1318
Db	1252	AGCCTCTGCTCTCCAGAAATCTAAGCCTCTTACAGAACTTAAGCAATATGATTCAGCCAT	1311
Qy	1319	GCTTGGAGGCTCTCTCAGCTTGGACCGTCTGCTTGCCTCCCGGGAGCCAGGGGGCCCC	1378
Db	1312	GTTGGAGAACTCTCTGATGTGAACGTTTGTGGGTTTCCAGGGAAGCCAAAGGAGCCCC	1371
Qy	1379	TCTGTTTCAGTACCCCAAGCGAGCGGATGGTGTCTAATGAAGACAGTAGAAGAGGA	1438
Db	1372	CCTACTGAATACCCCAAGCGAGAACGAATGGTGTCTATGAGACATGGAGGAAAGAA	1431
Qy	1439	CCTAGAGATTGAGAGGCTTAAAGACGAAGCAAAAGAAATCTGAGGCCAAGATGTTGGCCCCA	1498
Db	1432	CTTAGAGATTGAGAGGCTTAAAGATGAAGCAAAAGAAATCTGGAAGCCAAAGGTTGTTGGCCCCA	1491
Qy	1499	GAAGGCTGAGGA	1510
Db	1492	GGAGGCCACCGGA	1503

RESULT 6	1124 bp	mrna	linear	EST 25-MAR-2004
AL545712				
LOCUS				
DEFINITION	AL545712 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CS0DI009YJ01 5-PRIME, mrna sequence.			
ACCESSION	AL545712			
VERSION	AL545712.3	GI:45746191		

KEYWORDS	REFERENCE
SOURCE	AUTHOR
ORGANIZATION	TITLE
	JOURNAL
	COMMENT

EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 1124)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced qi:31267547.

FEATURES

source	1. .1124
Location/Qualifiers	
For more information about this cluster, see http://www.genoscope.cns.fr/cdia?e=CSODI009C801QP1&c=7766.r .	

FEATURES

ORIGIN

Query Match	68.6%;	Score 1054.4;	DB 1;	Length 1124;
Best Local Similarity	96.5%;	Pred. No. 3.9e-268;		
Matches 1068;	Conservative 25;	Mismatches 12;	Indels 2;	Gaps 2;
Qy	20	GAGCGACGAGCGACGCGAGATGCGCAGCTTCAGCGGGCGGCATCTCAGGAGCTGG	79	
Db	1	GAGCGACGAGCGACGCGAGATGCGGAGTTACGGGCGGCATCTCAGGAGCTGG	60	
Qy	80	TCGCTCTCGGCTAAGCAAGATTGGAGCTACTCGTCTGTCCACCTCCAGCTCGCCTAAGGGT	139	
Db	61	TCGCTGTGGCTAAGCAAGATTGGAGCTACTCGTCTGTCCACCTCCAGCTCGCTAAGGGT	120	
Qy	140	GGCTGTGCGACTGCGGCCCATTTGTGATGGAACAGCGGAGCAAGTGATCCCCCTGTGT	199	
Db	121	GGCTGTGCGACTGCGGCCCATTTGTGATGGAACAGCGGAGCAAGTGATCCCCCTGTGT	180	
Qy	200	GCGGGGCATGGACAGCTGCTCTTAGAGATTGCTAACTGGAGGAACACCGAGAGACTCT	259	
Db	181	GCGGGGCATGGACAGCTGCTCTTAGAGATTGCTAACTGGAGGAACACCGAGAGACTCT	240	
Qy	260	CAAAATACCAAGTTTGATGTCCTTTATGGGAGAGGAGTACTCAGCAGGACATCTATGCAGG	319	
Db	241	CAAAATACCAAGTTTGATGTCCTTTATGGGAGAGGAGTACTCAGCAGGACATCTATGCAGG	300	
Qy	320	TTCAAGTGACGCCCATCCTAAGGCATCTTGGAAGGGCAGAAATGCCAGTGTGTTGCCCTA	379	
Db	301	TTCAAGTGACGCCCATCCTAAGGCATCTTGGAAGGGCAGAAATGCCAGTGTGTTGCCCTA	360	
Qy	380	TGGACCCACAGGAGCTGGGAAGACGACACAATGCTTGGGCGCCACAGAGCAACTGGGT	439	
Db	361	TGGACCCACAGGAGCTGGGAGACGACACAATGCTTGGGAGCCACAGAGCAACTGGGGK	420	
Qy	440	GATCCCGCGGGTCTCATGGAACCTCTCGACGCTCAACAAGGAGAGGGGTGCCAGGGCCG	499	
Db	421	GATCCCGCGGGTCTCATGGAACCTCTCGACGCTCAACAAGGAGAGGGGTGCCAGGGSCG	480	
Qy	500	GCATGGGGCCCTTTCTGTGCACCATGTCTTAACCTAGAGATCTACCGAGAGAGGTATTAGA	559	

```
Db 481 GCCATGGCCCTTCTCTCACCATTCTTACCTAGAGATCTACCGAGGAGAGKATAGA 540
Qy 560 CTTCTCGACCTGCTTCGGGAGACCTGGTAATCCGAGAGAGCTGCCGGGGAATATCCT 619
Db 541 CTTCTCGACCTGCTTCGGGAGACCTGGTAATCCGAGAGAGCTGCCGGGGRATATCCT 600
Qy 620 GATTCCGGGTCTCTCCAGAGGCCATCAGTAGCTTTGCTGATTTGAGCGGCACCTCCT 679
Db 601 GATTCCGGGTCTCTCCAGAGGCCATCAGTAGCTTTGCTGATTTGAGCGGCACCTCCT 660
Qy 680 GCCAGGCAGTCGAATCCGAGCTGTAGGAGCACCCGGCTCAACCCAGGCTCTCCCGCAG 739
Db 661 GCCAGGCAGTCGAATCCGAGCTGTAGGAGCACCCGGCTCAACCCAGGCTCTCCCGCAG 720
Qy 740 TCATGCTGCTCTCTGCTCAAGGTGGACACCGGAGACCTTTGGCCCCCAATTCGCCAGCG 799
Db 721 TCATGCTGCTCTCTGCTCAAGGTGGACACCGGAGACCTTTGGCCCCCAATTCGCCAGCG 780
Qy 800 AGAGGGAAACTTACTGATTGACTTGGCTGGGTGAGGAGCAACCGGCGGACAGGCAA 859
Db 781 AGAGGGAAACTTACTGATTGACTTGGCTGGGTGAGGAGCAACCGGCGGACAGGCAA 840
Qy 860 CAAGGGCTCTCGGT-AAAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCGGCA 918
Db 841 CAAGGGCTCTCGSTAAAAAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCGGCA 900
Qy 919 AAGTGTAGATGCGTGAATCAGGGCTCCTCGTGTACTTATCGGACAGCAAGCTCA 978
Db 901 AAGTGTAGATGCGTGAATCAGGGCTCCTCGTGTACTTATCGGACAGCAAGCTCA 960
Qy 979 CTGCGCTATTGAGGACTCTCGGTGGCTCAGCCACAGTATCCTTATGCGCAACATTG 1038
Db 961 CTGCGCTATTGAGGACTCTCTGGGKGGSTCAG-CCACAGTATCCTTATGCGCAACATTG 1019
Qy 1039 CCCTCGAGAGACGCTTCTACTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGTCCA 1098
Db 1020 CCCTCGAGAGACGCTTCTACTAGACACAGTCTCCGCACTCAACTTTGCTGSGAGKCCA 1079
Qy 1099 AGAGGTGATCAATCGCGCTTTTACCA 1125
Db 1080 AGGGGKGRKTAATCGGCTTTTANCAA 1106

RESULT 7
BX337406
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1027)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 2, 2003 this sequence version replaced gi:30341614.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7766.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DI041AB07QP1&c=7766.r.
```

```
FEATURES
source
Location/Qualifiers
1..1027
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI041Y113"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 65.2%; Score 1002.8; DB 5; Length 1027;
Best Local Similarity 99.5%; Pred. No. 1.8e-254;
Matches 1023; Conservative 3; Mismatches 0; Indels 2; Gaps 2;
Qy 38 CGAGATGGCGGAGCTTCCAGC-GGGCGGCGATCTCAGGAGCTGCTCGCTCGGCTAAGCA 96
Db 1 CGAGATGGCGGAGCTTCCAGCGGGCGGCGATCTCAGGAGCTGCTCGCTCGGCTAAGCA 60
Qy 97 AGATTGGAGCTACTCGTCTCCACCTCCAGCTCCGCTAAAGGCTGGCTGTCCGACTCGGC 156
Db 61 AGATTGGAGCTACTCGTCTCCACCTCCAGCTCCGCTAAAGGCTGGCTGTCCGACTCGGC 120
Qy 157 CATTTGTGATGGAACAGCGGGAGCAAGTGTATCCCTCTGTGTGCGGGGATGGAACAGCT 216
Db 121 CATTTGTGATGGAACAGCGGGAGCMAGTGTATCCCTCTGTGTGCGGGGATGGAACAGCT 180
Qy 217 GCTCTCTAGAGATGCTTAATCTGAGGAAACACAGGAGACTCTCAAAATACCAAGTTGATG 276
Db 181 GCTCTCTAGAGATGCTTAATCTGAGGAAACACAGGAGACTCTCAAAATACCAAGTTGATG 240
Qy 277 CTTTCTATGGGAGAGAGTACTCAGCAGGACATCTATCAGAGTTTCAGTGCAGCCCATCC 336
Db 241 CTTTCTATGGGAGAGAGTACTCAGCAGGACATCTATCAGAGTTTCAGTGCAGCCCATCC 300
Qy 337 TAAGGACATTGCTGGAAGGGCAGATGCCAGTGTGCTTGCCTATGAGCCACAGAGAGCTG 396
Db 301 TAAGGACATTGCTGGAAGGGCAGATGCCAGTGTGCTTGCCTATGAGCCACAGAGAGCTG 360
Qy 397 GGAAGACGACACAAATGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 456
Db 361 GGAAGACGACACAAATGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Qy 457 TGGACCTCTCGAGCTCACAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 516
Db 421 TGGACCTCTCGAGCTCACAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 479
Qy 517 TCACCATGTTTACCTAGAGATCTACGAGAGAGAGTATAGACCTCTCGGACCTGCTT 576
Db 480 TCACCATGTTTACCTAGAGATCTACGAGAGAGAGTATAGACCTCTCGGACCTGCTT 539
Qy 577 CGGAGACCTGTTAATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 636
Db 540 CGGAGACCTGTTAATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 599
Qy 637 AGAAGCCCATCAGTAGCTTTGCTGATTTTGAAGCGGACCTTCTTCGAGCAGGAGTCAAAATC 696
Db 600 AGAAGCCCATCAGTAGCTTTGCTGATTTTGAAGCGGACCTTCTTCGAGCAGGAGTCAAAATC 659
Qy 697 GGAAGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 756
Db 660 GGAAGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 719
Qy 757 TCAAGGTGAGACAGCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 816
Db 720 TCAAGGTGAGACAGCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 779
Qy 817 TGATTGACTTGGCTGGGTGAGAGGACAAACCGGCGGACAGGCAACAGGAGGCTTTGGGCTAA 876
Db 780 TGATTGACTTGGCTGGGTGAGAGGACAAACCGGCGGACAGGCAACAGGAGGCTTTGGGCTAA 839
```



```

VERSION BX37353.2 GI:46272059
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 1032)
JOURNAL Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
COMMENT Full-length cDNA libraries and normalization
Unpublished (2001)
On May 2, 2003 this sequence version replaced gi:30337600.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7766.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DK01040B030P1&c=7766.r.
FEATURES
Location/Qualifiers
1..1032
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK01040Y106"
/tissue types="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 63.6%; Score 978.6; DB 5; Length 1032;
Best Local Similarity 98.6%; Pred. No. 4.8e-248;
Matches 1003; Conservative 4; Mismatches 8; Indels 2; Gaps 2;
QY 70 CAGGAGCTGCTGCTGGCTAAGCAAGATTGGAGCTACTCTGCTCCACCTCCAGCTC 129
DB 1 CAGGAGCTGCTGCTGGCTAAGCAAGATTGGAGCTACTCTGCTCCACCTCCAGCTC 60
QY 130 GCGTAAGGGTGGCTGCGACTGCGGCAATTTGGATGAAACAGCGGGAGCAAGTGATC 189
DB 61 GCGTAAGGGTGGCTGCGACTGCGGCAATTTGGATGAAACAGCGGGAGCAAGTGATC 120
QY 190 CCCCCTGTGTGCGGGGATGGAGAGCTGCTCTAGAGATTGCTAACTGGAGAACCCACC 249
DB 121 CCCCCTGTGTGCGGGGATGGAGAGCTGCTCTAGAGATTGCTAACTGGAGAACCCACC 180
QY 250 AGGAGACTCTCAATACCAATACCAATTTGATGCTTATGGGAGAGAGTACTCAGCAGACA 309
DB 181 AGGAGACTCTCAATACCAATTTGATGCTTATGGGAGAGAGTACTCAGCAGACA 240
QY 310 TCTATGAGGTTGAGTGCAGCCATCTTAAGGCACTTGTGGAAGGGCAGAAATGCCAGTG 369
DB 241 TCTATGAGGTTGAGTGCAGCCATCTTAAGGCACTTGTGGAAGGGCAGAAATGCCAGTG 300
QY 370 TGCTTGCCTATGACCCACAGGAGCTGGGAGAGCGACAACTGCTGGGAGCCAGAGC 429
DB 301 TGCTTGCCTATGACCCACAGGAGCTGGGAGAGCGACAACTGCTGGGAGCCAGAGC 360
QY 430 AACCTGGGTGATCCCGCGGGCTCTCATGGACCTCTCTGAGCTCAACAAGGAGAGGGTG 489
DB 361 AACCTGGGTGATCCCGCGGGCTCTCATGGACCTCTCTGAGCTCAACAAGGAGAGGGTG 420
QY 490 CCGAGGGCCGGCCATGGGCCCTTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGA 549
DB 421 CCGAGGGCCGGCCATGGGCCCTTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGA 480

```

```

QY 550 AGGTATTAGACCTCTGGACCTCTTGGGAGACCTGGTAAATCCGAGAGAGATGCGGG 609
DB 481 AGGTATTAGACCTCTGGACCTCTTGGGAGACCTGGTAAATCCGAGAGAGATGCGGG 540
QY 610 GGAATATCTGATTTCCGGGTCTCTCCAGAAAGCCCATCAGTAGCTTGTCTGATTTTGAGC 669
DB 541 GGAATATCTGATTTCCGGGTCTCTCCAGAAAGCCCATCAGTAGCTTGTCTGATTTTGAGC 600
QY 670 GGCATCTTCTGCGAGCAGTCGAAATCGGACTGTAGAGCCACCCGGCTCAACAGCGCT 729
DB 601 GGCATCTTCTGCGAGCAGTCGAAATCGGACTGTAGAGCCACCCGGCTCAACAGCGCT 660
QY 730 CTCTCCGCGAGTCATGCTGCTCTCTGCTCAAGTGGAGCCAGCGGGAACGTTTGGCCCAT 789
DB 661 CTCTCCGCGAGTCATGCTGCTCTCTGCTCAAGTGGAGCCAGCGGGAACGTTTGGCCCAT 720
QY 790 TTCCGCGAGGAGAGGGAACCTCTACTGATTTGCTTGGCTGGGTGAGAGCAACCGGC 849
DB 721 TTCCGCGAGGAGAGGGAACCTCTACTGATTTGCTTGGCTGGGTGAGAGCAACCGGC 780
QY 850 GCACAGGCAACAGGGCTTGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTG 909
DB 781 GCACAGGCAACAGGGCTTGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTG 840
QY 910 TCCTGGGCAAGTGTAGATGCGCTGAATCAGGGCTCCCTCGT-GTACCTTATCGGAC 968
DB 841 TCCTGGGCAAGTGTAGATGCGCTGAATCAGGGCTCCCTCGTGGTACCTTATCGGAC 900
QY 969 AGCAAGCTCACTGCTCTTATTCAGGACTCTC-TGGGTGGCTCAGGCCACAGTATCTCTAT 1027
DB 901 AGCAAGCTCACTGCTCTTATTCAGGACTCTCTTGGGTGGSTCAGGCCACATGATCTTAT 960
QY 1028 TGCCAACTTGGCCCTGAGAGAGCTTCTTACCTAGACAGAGTCTCCGACTCAACTT 1084
DB 961 TGCCAACTTGGCCCTGAGAGAGCTTCTTATCTAGACAGAGTCTGGSCTCACTT 1017

```

```

RESULT 10
BX363358 959 bp mRNA linear EST 08-APR-2004
LOCUS BX363358 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION CDNA clone CS0DK012YE23 5-PRIME, mRNA sequence.
ACCESSION BX363358
VERSION BX363358.2 GI:46287820
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 959)
JOURNAL Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
COMMENT Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:30368590.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7766.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DK012AC12QPl&c=7766.r.
FEATURES
Location/Qualifiers
1..959
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK012YE23"
/cell_type="HELA CELLS COT 25-NORMALIZED"

```

```
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

Query Match 61.2%; Score 942; DB 5; Length 959;
Best Local Similarity 99.9%; Pred. No. 2.4e-238;
Matches 953; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 6 GCCGGCGCGCTCAGCAGCAGAGCGAGCGCGAGATGGCGGAGCTTCCAGCGG-CGGC 64
DB 6 GCCGGCGCGCTCAGCAGCAGAGCGAGCGCGAGATGGCGGAGCTTCCAGCGGCGGC 65
QY 65 GATCTCAGAGCTGTGCTCGCTAAGCAAGATTGGAGCTACTGCTCTCACTCC 124
DB 66 GATCTCAGAGCTGTGCTCGCTAAGCAAGATTGGAGCTACTGCTCTCACTCC 125
QY 125 AGCTCGCGTAAGGCTGTGCTCGCTAAGCAAGATTGGAGCTACTGCTCTCACTCC 184
DB 126 AGCTCGCGTAAGGCTGTGCTCGCTAAGCAAGATTGGAGCTACTGCTCTCACTCC 185
QY 185 TGATCCCCCTGTGTGCGGGATGGAAGCTGTCTCTAGAGATTGCTAACTGGAGAA 244
DB 186 TGATCCCCCTGTGTGCGGGATGGAAGCTGTCTCTAGAGATTGCTAACTGGAGAA 245
QY 245 CCACGAGAGACTCTCAATACCAAGTTGATGCTCTCTATGGGAGAGGAGTACTCAGCA 304
DB 246 CCACGAGAGACTCTCAATACCAAGTTGATGCTCTCTATGGGAGAGGAGTACTCAGCA 305
QY 305 GGACATCTATGCAAGTTCAGTGCAAGCCATCTTAAGGCACCTTCTGGAAGGCGAGAATGC 364
DB 306 GGACATCTATGCAAGTTCAGTGCAAGCCATCTTAAGGCACCTTCTGGAAGGCGAGAATGC 365
QY 365 CAGTGTGCTTCCATGAGACCCACAGGAGCTGGGAAGACGACACAACTGTGGGCGAGCCC 424
DB 366 CAGTGTGCTTCCATGAGACCCACAGGAGCTGGGAAGACGACACAACTGTGGGCGAGCCC 425
QY 425 AGACCACTTGGGTGATCCCGGGCTCTCATGACCTCTGACGCTCAGAGGAGGA 484
DB 426 AGACCACTTGGGTGATCCCGGGCTCTCATGACCTCTGACGCTCAGAGGAGGA 485
QY 485 GGGTCCGAGGCGCGGACATGGGCTTCTGTCACCATGCTTACCTAGAGATCTACCA 544
DB 486 GGGTCCGAGGCGCGGACATGGGCTTCTGTCACCATGCTTACCTAGAGATCTACCA 545
QY 545 GGAGAAAGTATTAGACTCTTGGACCCCTGCTTCGGGAGACCTGTTAATCCGAGAGACTG 604
DB 546 GGAGAAAGTATTAGACTCTTGGACCCCTGCTTCGGGAGACCTGTTAATCCGAGAGACTG 605
QY 605 CCGGGGGAATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTT 664
DB 606 CCGGGGGAATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTT 665
QY 665 TGAGCGGCACTTCTGCGGAGCAGTGGAAATCGGACTGTAGAGCCACCGGGCTCAACCA 724
DB 666 TGAGCGGCACTTCTGCGGAGCAGTGGAAATCGGACTGTAGAGCCACCGGGCTCAACCA 725
QY 725 GCGTCTCTCCGAGTCACTGCTGCTCTGCTCAAGGTGGACGAGGGAACGTTTGGC 784
DB 726 GCGTCTCTCCGAGTCACTGCTGCTCTGCTCAAGGTGGACGAGGGAACGTTTGGC 785
QY 785 CCAATTTGCGCAGGAGGGAATACTTCACTGATTTGA CTGCTGGGTGAGAGGACAA 844
DB 786 CCAATTTGCGCAGGAGGGAATACTTCACTGATTTGA CTGCTGGGTGAGAGGACAA 845
QY 845 CCGGCGCACAGGCAAGGSCCTTCGCTCTAAAGAGGTGAGGCCATCAACCTCCCT 904
DB 846 CCGGCGCACAGGCAAGGSCCTTCGCTCTAAAGAGGTGAGGCCATCAACCTCCCT 905
QY 905 GTTTGTCTGGGCAAGTGTGATGCGCTGAAATCAGGCGCCTCCCTCGTGATCC 958

DB 906 GTTTGTCTGGGCAAGTGTGATGCGCTGAAATCAGGCGCCTCCCTCGTGATCC 959

RESULT 11
LOCUS BX354166
DEFINITION BX354166 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DC015VH04 5-PRIME, mRNA sequence.
ACCESSION BX354166
VERSION BX354166.2 GI:46304510
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 958)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 5, 2003 this sequence version replaced gi:30373827.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7766.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DC015DD02QPI&c=7766.r.

FEATURES
source
1..958
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC015VH04"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 60.7%; Score 934; DB 5; Length 958;
Best Local Similarity 99.8%; Pred. No. 3.2e-236;
Matches 956; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 7 CCGCGGCGGCTCGACCGCAGCAGGCGGACGCGAGATGGCGGAGCTTACGGC-CGGC 65
DB 1 CCGCGGCGGCTCGACCGCAGCAGGCGGACGCGAGATGGCGGAGCTTACGGC-CGGC 60
QY 66 ATCTCAGAGCTGTGCTCGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCA 125
DB 61 ATCTCAGAGCTGTGCTCGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCA 120
QY 126 GCTCGCTAAGGCTGTGCTCGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCA 185
DB 121 GCTCGCTAAGGCTGTGCTCGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCA 180
QY 186 GATCCCCCTGTGTCGGGCGATGGACAGCTGCTCTAGAGATTGCTAACTGAGGAGAAC 245
DB 181 GATCCCCCTGTGTCGGGCGATGGACAGCTGCTCTAGAGATTGCTAACTGAGGAGAAC 240
QY 246 CACGAGAGACTCTCAATACCAAGTTGATGCTCTTATGGGAGAGGAGTACTCAGCAG 305
DB 241 CACGAGAGACTCTCAATACCAAGTTGATGCTCTTATGGGAGAGGAGTACTCAGCAG 300
QY 306 GACATCTATGCAAGTTCAGTGCAAGCCATCTTAAGGCACCTTCTGGAAGGCGAGAATGCC 365

[illegible]

RESULT 13	AL559769	LOCUS	AL559769	999 bp	mRNA	linear	EST 02-APR-2004
DEFINITION	AL559769 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens CDNA clone CS0DG005YF11 5-PRIME, mRNA sequence.						

EST. 2376823 Q1:40205441

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 999)

Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished (2001)

JOURNAL

COMMENT On Feb 15, 2001 this sequence version replaced gi:31283900.

```

FEATURES
source
1. .999
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DG005Yf11"
/tissue_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

```

LOCUS	1094 bp	linear	EST 27-APR-2004
DEFINITION	BX376199 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens		
DEFINITION	cDNA clone CS0DC024YA10 5-PRIME, mRNA sequence.		
ACCESSION	BX376199		
VERSION	BX376199.2		
KEYWORDS	GI:46616791		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
COMMENT	1 (bases 1 to 1094)		
COMMENT	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.		
COMMENT	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
COMMENT	On May 8, 2003 this sequence version replaced gi:30452476.		
COMMENT	Contact: Genoscope		
COMMENT	Genoscope - Centre National de Sequencage		
COMMENT	BP 191 91006 EVRY cedex - France		

Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7766.1

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?s=CS0DC024BA05QP1&c=77666.r>

FEATURES

source

```

l. 1094
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cclone="CS0DC024YA10"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/cclone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

	Query Match	59.9%; Score 921.4; DB 5; Length 1094;	
	Best Local Similarity	99.2%; Pred. No. 7.2e-233;	
	Matches 963; Conservative 4; Mismatches 0; Indels 4; Gaps 4;		
QY	6	GCCGCGCGCGCTCGACGACAGCAGCGAGATGGCGGAGCTTTCAGCGG-CGGC	64
Db	12	: :	71
QY	65	GATCTCAGGAGCTGTGCCTGTCTGGCTTAAGCAAGATTGGAGCTACTCGTCGTCCACCTCC	124
Db	72	GATCTCAGGAGCTGGTCGCTGTGGCTAAGCAAGATTGGAGCTACTCGTCGTCCACCTCC	131
QY	125	AGCTCGGTTAAGGTTGGCTGTGCGACTGCGGCCATTGTTGTGATGGAAACAGCGGGAGCGAG	184
Db	132	AGCTCGGTTAAGGTTGGCTGTGCGACTGCGGCCATTGTTGTGATGGAAACAGCGGGAGCGAG	191
QY	185	TGATCCCCCTGTGTGGGGGCGCATGGACAGCTGTCTCTAGAGATTGTCTAACTCGAGGAA	244
Db	192	TGATCCCCCTGTGTGGGGGCGCATGGACAGCTGTCTCTAGAGATTGTCTAACTCGAGGAA	251
QY	245	CACACGAGGACTCTCAAATAACAGTTTGATGCTCTTATATGGGAGAGGAGTACTCAGCA	304
Db	252	CCACGAGGACTCTCAAATAACAGTTTGATGCTCTTATATGGGAGAGGAGTACTCAGCA	311
QY	305	GGACATCTATGCAGGTTTCAGTGACAGCCCATCTTAAGGCACCTTGCTGGAAGGGCAGAATGC	364
Db	312	GGACATCTATGCAGGTTTCAGTGACAGCCCATCTTAAGGCACCTTGCTGGAAGGGCAGAATGC	371
QY	365	CAGTGTGCTTGCCTATGACCCTCACAGAGCTGGGAAGACGCACAATGCTGGGCAGCCC	424
Db	372	CAGTGTGCTTGCCTATGACCCTCACAGAGCTGGGAAGACGCACAATGCTGGGCAGCCC	431
QY	425	AGAGCAACCTGGGGTGAATCCGGGGGTCTCATGGACCTCTCTGAGCTCACAGGGAGGA	484
Pb	432	AGAGCAACCTGGGGTGAATCCGGGGGTCTCATGGACCTCTCTGAGCTCACAGGGAGGA	491
QY	485	GGGTGCGAGGGCGCGGCATGGGCCCCTTCTGTCAACCATGTCTTACCTAGAGATCTACCA	544
Db	492	GGGTGCGAGGGCGCGGCATGGGCCCCTTCTGTCAACCATGTCTTACCTAGAGATCTACCA	551
QY	545	GGAGAGGTAATTAGACCTCTCTGAGACCTGCTTGGGAGACTGTGGTAATCCGAGAAGACTG	604
Db	552	GGAGAGGTAATTAGACCTCTCTGAGACCTGCTTGGGAGACTGTGGTAATCCGAGAAGACTG	611
QY	605	CCGGGGGAATATCTGAATCCGGGTCTCTCCAGAACCCCATCAGTAGCTTTGCTGATTT	664
Db	612	CCGGGGGAATATCTGAATCCGGGTCTCTCCAGAACCCCATCAGTAGCTTTGCTGATTT	671
QY	665	TGAGCGGCACTTCTCTGCCAGCCAGTGGAAATCGGACTGTGAGGAGCCACCCGGGTCAAACA	724

ORIGIN

Query Match 59.1%; Score 909; DB 4; Length 1019;
Best Local Similarity 97.5%; Pred. No. 1.4e-229;
Matches 965; Conservative 0; Mismatches 21; Indels 4; Gaps 4;

Db	672	TGAGCGGCACTTCTCGCAGCCAGTCGGAATTCGAGCTGTAGGAGCCACCCGGCTCAACCA	731
Qy	725	CGCTCTCCCGCAGTCATGCTGTGCTCTCGTCAAGGTGACACAGCGGGAAACGTTTGGC	784
Db	732	CGCTCTCCCGCAGTCATGCTGTGCTCTCGTCAAGGTGACACAGCGGGAAACGTTTGGC	791
Qy	785	CCGATTTTCGCAGCGAGAGGGAAAACCTACCTGATTGCATTTGGCTGGGTGAGGACAA	844
Db	792	CCCATTTTCGCAGCGAGAGGGAAAACCTACCTGATTGCATTTGGCTGGGTGAGGACAA	851
Qy	845	CCGCGCACAGGCAACAAAGGCTTCCTGGCT-AAAAGAGAGTGGAGCCATCAACACCTCC	903
Db	852	CCGCGCACAGGCAACAAAGGGSCTTCGCTAAAAAGAGTGGAGCCATCAACACCTCC	911
Qy	904	TGTTTGTCTGGGCAAGTGGTAGATGGCGTGAATCAGGGGCTCCCTGGTGTACTTATC	963
Db	912	TGTTTGTCTGGGCAAGTGGTAGATGGCGTGAATCAGGG-CTCCCTCGTGTA-CTTTATC	969
Qy	964	GGGACGACAAG	974
Db	970	GGGACGACAAS	980

RESULT 15
BM546877

BM546877	1019 bp	mrna	linear	EST 20-FEB-2002
LOCUS				
DEFINITION	AGENCOURT_6491037 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5723623			
	5', mRNA sequence.			
ACCESSION	BM546877			
VERSION	BM546877.1			
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
TITLE	1 (bases 1 to 1019)			
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .			
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)			
	Unpublished (1999)			
	Contact: Robert Strausberg, Ph.D.			
	Email: cgabs-r@mail.nih.gov			
	Tissue Procurement: Invitrogen			
	cDNA Library Preparation: Life Technologies, Inc.			
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
	DNA Sequencing by: Agencourt Bioscience Corporation			
	Clone distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LLNL at:			
	http://image.llnl.gov			
	Plate: LLAM12711	row: j	column: 08	
	High quality sequence stop: 688.			

FEATURES

SOURCES

```

1: .roid3
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5723623"
/lab_host="DH10B"
/clone_lib="NIH_MGC_135"
/notes="Ecorv: ovary (pool of 3); Vector: pCMV-SPORT6;
Site.1: Ecorv (destroyed); Site.2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(Ecorv site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

```


Qy 21 ACGCAGCAGAGCGGCGACGCGAGATGCGCGCAGCTTCAGCGCGCGGCGATCTCAGGAGCTGGT 80
Db |||||
1 ACGCAGCAGAGCGGCGACGCGAGATGCGCGCAGCTTCAGCGCGCGGCGATCTCAGGAGCTGGT 60
Qy 81 CGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGGTCCACCTCCAGCTCGCGTAAAGGGTG 140
Db |||||
61 CGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGGTCCACCTCCAGCTCGCGTAAAGGGTG 120
Qy 141 GCTGTGGACCTGCGGCCATTGTGTGATGGAACAGCGGGAGCAAGTGATGCCCCCTGTGTG 200
Db |||||
121 GCTGTGGACCTGCGGCCATTGTGTGATGGAACAGCGGGAGCAAGTGATGCCCCCTGTGTG 180
Qy 201 CGGGGCATGGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACTCTC 260
Db |||||
181 CGGGGCATGGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACTCTC 240
Qy 261 AAATACCAAGTTTGATGCCCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCAGGT 320
Db |||||
241 AAATACCAAGTTTGATGCCCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCAGGT 300
Qy 321 TCAGTGCAGCCCATCTTAAGGCACCTTGCTGGAAGGGCAGAAATGCCAGTGTGCTTGCCTAT 380
Db |||||
301 TCAGTGCAGCCCATCTTAAGGCACCTTGCTGGAAGGGCAGAAATGCCAGTGTGCTTGCCTAT 360
Qy 381 GGACCCACAGAGCTGGGAAGACGACACAAATGCTGGGCGAGCCAGAGCAACCTGGGGTG 440
Db |||||
361 GGACCCACAGAGCTGGGAAGACGACACAAATGCTGGGCGAGCCAGAGCAACCTGGGGTG 420
Qy 441 ATCCGGGGGCTCTATGGACCTCTGTCAGCTCACAAGGGAGGAGGGTGCCGAGGGCCGG 500
Db |||||
421 ATCCGGGGGCTCTCTATGGACCTCTGTCAGCTCACAAGGGAGGAGGGTGCCGAGGGCCGG 480
Qy 501 CCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGGAGGATATTAGAC 560
Db |||||
481 CCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGGAGGATATTAGAC 540
Qy 561 CTCTGGACCTGCTGGGAGACCTGGTAATCCGAGAGACTGCGGGGGAAATATCCTG 620
Db |||||
541 CTCTGGACCTGCTGGGAGACCTGGTAATCCGAGAGACTGCGGGGGAAATATCCTG 600
Qy 621 ATTGGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGCGGCACTTCTG 680
Db |||||
601 ATTGGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGCGGCACTTCTG 660
Qy 681 CCAGCCAGTGGAAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCCCGGAGT 740
Db |||||
661 CCAGCCAGTGGAAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCCCGGAGT 720
Qy 741 CATGCTGTCTCTGGTCAAGGTGGACCCAGCGGAAAGTTTGG-CCCCATTTCCGCCAGCG 799
Db |||||
721 CATGCTGTCTCTGGTCAAGGTGGACCCAGCGGAAAGTTTGG-CCCCATTTCCGCCAGCG 780
Qy 800 AGAGGGAAACTTACCTGATTTGCTGGGTGAGAGCAACCCGGCGCAGAGCAA 859
Db |||||
781 AGAGGGAAACTTACCTGATTTGCTGGGTGAGAGCAACCCGGCGCAGAGCAA 840
Qy 860 CAAGGGCTTCGGCTAAAGAGTGGAGCCATCAACACCTCCCTGTTGCTTGGGCAA 919
Db |||||
841 CAAGGGCTTCGGCTAAAGAGTGGAGCCATCAACACCTCCCTGTTGCTTGGGCAA 900
Qy 920 AGTGGTAGCTGCTGAATCAGGGCCCTCCCTGCTGTA-CTTTATCGGAGCAGCAA-GCTC 977
Db |||||
901 AGTGGTAGCTGCTGAATCAGGGCCCTCCCTGCTGTA-CTTTATCGGAGCAGCAA-GCTC 960
Qy 978 ACTGCTATTGCA-GGACTCTGGGTGG 1006
Db |||||
961 ACTGCTATTGCAAGGACTCTCTGGGTGG 990

This Page Blank (uspto)